

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
7 June 2001 (07.06.2001)

PCT

(10) International Publication Number
WO 01/40269 A2

(51) International Patent Classification⁷: C07K 14/00

Jennifer, L. [US/US]; 16677 NE 88th Street, Redmond, WA 98052 (US). WANG, Aijun [CN/US]; 3106 213th Place SE, Issaquah, WA 98029 (US).

(21) International Application Number: PCT/US00/32520

(74) Agents: POTTER, Jane, E., R.; Seed Intellectual Property Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7092 et al. (US).

(22) International Filing Date: 29 November 2000 (29.11.2000)

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CI, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(25) Filing Language: English

(26) Publication Language: English

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(30) Priority Data:

09/451,651	30 November 1999 (30.11.1999)	US
09/510,662	22 February 2000 (22.02.2000)	US
09/523,586	10 March 2000 (10.03.2000)	US
09/545,068	7 April 2000 (07.04.2000)	US
09/571,025	15 May 2000 (15.05.2000)	US

(71) Applicant (for all designated States except US): CORIXA CORPORATION [US/US]; Suite 200, 1124 Columbia Street, Seattle, WA 98104 (US).

Published:

— Without international search report and to be republished upon receipt of that report.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



A2

WO 01/40269

(54) Title: COMPOSITIONS AND METHODS FOR THERAPY AND DIAGNOSIS OF BREAST CANCER

(57) Abstract: Compositions and methods for the therapy and diagnosis of cancer, such as breast cancer, are disclosed. Compositions may comprise one or more breast tumor proteins, immunogenic portions thereof, or polynucleotides that encode such portions. Alternatively, a therapeutic composition may comprise an antigen presenting cell that expresses a breast tumor protein, or a T cell that is specific for cells expressing such a protein. Such compositions may be used, for example, for the prevention and treatment of diseases such as breast cancer. Diagnostic methods based on detecting a breast tumor protein, or mRNA encoding such a protein, in a sample are also provided.

COMPOSITIONS AND METHODS FOR THERAPY AND DIAGNOSIS
OF BREAST CANCER

TECHNICAL FIELD

The present invention relates generally to therapy and diagnosis of cancer, such as breast cancer. The invention is more specifically related to polypeptides comprising at least a portion of a breast tumor protein, and to polynucleotides encoding such polypeptides. Such polypeptides and polynucleotides may be used in vaccines and pharmaceutical compositions for prevention and treatment of breast cancer, and for the diagnosis and monitoring of such cancers.

10 BACKGROUND OF THE INVENTION

Breast cancer is a significant health problem for women in the United States and throughout the world. Although advances have been made in detection and treatment of the disease, breast cancer remains the second leading cause of cancer-related deaths in women, affecting more than 180,000 women in the United States each year. For women in North America, the life-time odds of getting breast cancer are now one in eight.

No vaccine or other universally successful method for the prevention or treatment of breast cancer is currently available. Management of the disease currently relies on a combination of early diagnosis (through routine breast screening procedures) and aggressive treatment, which may include one or more of a variety of treatments such as surgery, radiotherapy, chemotherapy and hormone therapy. The course of treatment for a particular breast cancer is often selected based on a variety of prognostic parameters, including an analysis of specific tumor markers. *See, e.g.*, Porter-Jordan and Lippman, *Breast Cancer* 8:73-100 (1994). However, the use of established markers often leads to a result that is difficult to interpret, and the high mortality observed in breast cancer patients indicates that improvements are needed in the treatment, diagnosis and prevention of the disease.

Accordingly, there is a need in the art for improved methods for therapy and diagnosis of breast cancer. The present invention fulfills these needs and further provides other related advantages.

SUMMARY OF THE INVENTION

5 Briefly stated, the present invention provides compositions and methods for the diagnosis and therapy of cancer, such as breast cancer. In one aspect, the present invention provides polypeptides comprising at least a portion of a breast tumor protein, or a variant thereof. Certain portions and other variants are immunogenic, such that the ability of the variant to react with antigen-specific antisera is not substantially 10 diminished. Within certain embodiments, the polypeptide comprises a sequence that is encoded by a polynucleotide sequence selected from the group consisting of: (a) sequences recited in SEQ ID NO: 1-38, 42-204, 205, 207 and 210-290; (b) variants of a sequence recited in SEQ ID NO: 1-38, 42-204, 205, 207 and 210-290; and (c) complements of a sequence of (a) or (b).

15 The present invention further provides polynucleotides that encode a polypeptide as described above, or a portion thereof (such as a portion encoding at least 15 amino acid residues of a breast tumor protein), expression vectors comprising such polynucleotides and host cells transformed or transfected with such expression vectors.

20 Within other aspects, the present invention provides pharmaceutical compositions comprising a polypeptide or polynucleotide as described above and a physiologically acceptable carrier.

25 Within a related aspect of the present invention, vaccines for prophylactic or therapeutic use are provided. Such vaccines comprise a polypeptide or polynucleotide as described above and an immunostimulant.

30 The present invention further provides pharmaceutical compositions that comprise: (a) an antibody or antigen-binding fragment thereof that specifically binds to a breast tumor protein; and (b) a physiologically acceptable carrier.

Within further aspects, the present invention provides pharmaceutical compositions comprising: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) a pharmaceutically acceptable carrier or excipient. Antigen

presenting cells include dendritic cells, macrophages, monocytes, fibroblasts and B cells.

Within related aspects, vaccines are provided that comprise: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) an 5 immunostimulant.

The present invention further provides, in other aspects, fusion proteins that comprise at least one polypeptide as described above, as well as polynucleotides encoding such fusion proteins.

Within related aspects, pharmaceutical compositions comprising a fusion 10 protein, or a polynucleotide encoding a fusion protein, in combination with a physiologically acceptable carrier are provided.

Vaccines are further provided, within other aspects, that comprise a fusion protein, or a polynucleotide encoding a fusion protein, in combination with an immunostimulant.

15 Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient a pharmaceutical composition or vaccine as recited above.

The present invention further provides, within other aspects, methods for removing tumor cells from a biological sample, comprising contacting a biological 20 sample with T cells that specifically react with a breast tumor protein, wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the protein from the sample.

Within related aspects, methods are provided for inhibiting the development of a cancer in a patient, comprising administering to a patient a biological 25 sample treated as described above.

Methods are further provided, within other aspects, for stimulating and/or expanding T cells specific for a breast tumor protein, comprising contacting T cells with one or more of: (i) a polypeptide as described above; (ii) a polynucleotide encoding such a polypeptide; and/or (iii) an antigen presenting cell that expresses such a 30 polypeptide; under conditions and for a time sufficient to permit the stimulation and/or

expansion of T cells. Isolated T cell populations comprising T cells prepared as described above are also provided.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a 5 patient an effective amount of a T cell population as described above.

The present invention further provides methods for inhibiting the development of a cancer in a patient, comprising the steps of: (a) incubating CD4⁺ and/or CD8⁺ T cells isolated from a patient with one or more of: (i) a polypeptide comprising at least an immunogenic portion of a breast tumor protein; (ii) a 10 polynucleotide encoding such a polypeptide; and (iii) an antigen-presenting cell that expresses such a polypeptide; and (b) administering to the patient an effective amount of the proliferated T cells, and thereby inhibiting the development of a cancer in the patient. Proliferated cells may, but need not, be cloned prior to administration to the patient.

15 Within further aspects, the present invention provides methods for determining the presence or absence of a cancer in a patient, comprising: (a) contacting a biological sample obtained from a patient with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; and (c) comparing the amount of polypeptide with a 20 predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within preferred embodiments, the binding agent is an antibody, more preferably a monoclonal antibody. The cancer may be breast cancer.

The present invention also provides, within other aspects, methods for monitoring the progression of a cancer in a patient. Such methods comprise the steps 25 of: (a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polypeptide detected in step (c) with the amount 30 detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

The present invention further provides, within other aspects, methods for determining the presence or absence of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a breast tumor protein; (b) detecting in the sample a level of a polynucleotide, preferably mRNA, that hybridizes to the oligonucleotide; and (c) comparing the level of polynucleotide that hybridizes to the oligonucleotide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within certain embodiments, the amount of mRNA is detected via polymerase chain reaction using, for example, at least one oligonucleotide primer that hybridizes to a polynucleotide encoding a polypeptide as recited above, or a complement of such a polynucleotide. Within other embodiments, the amount of mRNA is detected using a hybridization technique, employing an oligonucleotide probe that hybridizes to a polynucleotide that encodes a polypeptide as recited above, or a complement of such a polynucleotide.

In related aspects, methods are provided for monitoring the progression of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a breast tumor protein; (b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polynucleotide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

Within further aspects, the present invention provides antibodies, such as monoclonal antibodies, that bind to a polypeptide as described above, as well as diagnostic kits comprising such antibodies. Diagnostic kits comprising one or more oligonucleotide probes or primers as described above are also provided.

These and other aspects of the present invention will become apparent upon reference to the following detailed description. All references disclosed herein are hereby incorporated by reference in their entirety as if each was incorporated individually.

SEQUENCE IDENTIFIERS

SEQ ID NO: 1 is the determined cDNA sequence for clone 26915.
SEQ ID NO: 2 is the determined cDNA sequence for clone 26914.
SEQ ID NO: 3 is the determined cDNA sequence for clone 26673.
5 SEQ ID NO: 4 is the determined cDNA sequence for clone 26672.
SEQ ID NO: 5 is the determined cDNA sequence for clone 26671.
SEQ ID NO: 6 is the determined cDNA sequence for clone 26670.
SEQ ID NO: 7 is the determined cDNA sequence for clone 26669.
SEQ ID NO: 8 is a first determined cDNA sequence for clone 26668.
10 SEQ ID NO: 9 is a second determined cDNA sequence for clone 26668.
SEQ ID NO: 10 is the determined cDNA sequence for clone 26667.
SEQ ID NO: 11 is the determined cDNA sequence for clone 26666.
SEQ ID NO: 12 is the determined cDNA sequence for clone 26665.
SEQ ID NO: 13 is the determined cDNA sequence for clone 26664.
15 SEQ ID NO: 14 is the determined cDNA sequence for clone 26662.
SEQ ID NO: 15 is the determined cDNA sequence for clone 26661.
SEQ ID NO: 16 is the determined cDNA sequence for clone 26660.
SEQ ID NO: 17 is the determined cDNA sequence for clone 26603.
SEQ ID NO: 18 is the determined cDNA sequence for clone 26601.
20 SEQ ID NO: 19 is the determined cDNA sequence for clone 26600.
SEQ ID NO: 20 is the determined cDNA sequence for clone 26587.
SEQ ID NO: 21 is the determined cDNA sequence for clone 26586.
SEQ ID NO: 22 is the determined cDNA sequence for clone 26584.
SEQ ID NO: 23 is the determined cDNA sequence for clone 26583.
25 SEQ ID NO: 24 is the determined cDNA sequence for clone 26580.
SEQ ID NO: 25 is the determined cDNA sequence for clone 26579.
SEQ ID NO: 26 is the determined cDNA sequence for clone 26577.
SEQ ID NO: 27 is the determined cDNA sequence for clone 26575.
SEQ ID NO: 28 is the determined cDNA sequence for clone 26574.
30 SEQ ID NO: 29 is the determined cDNA sequence for clone 26573.
SEQ ID NO: 30 is the determined cDNA sequence for clone 25612.

SEQ ID NO: 31 is the determined cDNA sequence for clone 22295.
SEQ ID NO: 32 is the determined cDNA sequence for clone 22301.
SEQ ID NO: 33 is the determined cDNA sequence for clone 22298.
SEQ ID NO: 34 is the determined cDNA sequence for clone 22297.
5 SEQ ID NO: 35 is the determined cDNA sequence for clone 22303.
SEQ ID NO: 36 is the determined cDNA sequence for a first GABA_A receptor clone.
SEQ ID NO: 37 is the determined cDNA sequence for a second GABA_A receptor clone.
10 SEQ ID NO: 38 is the determined cDNA sequence for a third GABA_A receptor clone.
SEQ ID NO: 39 is the amino acid sequence encoded by SEQ ID NO: 36.
SEQ ID NO: 40 is the amino acid sequence encoded by SEQ ID NO: 37.
SEQ ID NO: 41 is the amino acid sequence encoded by SEQ ID NO: 38.
15 SEQ ID NO: 42 is the determined cDNA sequence for contig 1.
SEQ ID NO: 43 is the determined cDNA sequence for contig 2.
SEQ ID NO: 44 is the determined cDNA sequence for contig 3.
SEQ ID NO: 45 is the determined cDNA sequence for contig 4.
SEQ ID NO: 46 is the determined cDNA sequence for contig 5.
20 SEQ ID NO: 47 is the determined cDNA sequence for contig 6.
SEQ ID NO: 48 is the determined cDNA sequence for contig 7.
SEQ ID NO: 49 is the determined cDNA sequence for contig 8.
SEQ ID NO: 50 is the determined cDNA sequence for contig 9.
SEQ ID NO: 51 is the determined cDNA sequence for contig 10.
25 SEQ ID NO: 52 is the determined cDNA sequence for contig 11.
SEQ ID NO: 53 is the determined cDNA sequence for contig 12.
SEQ ID NO: 54 is the determined cDNA sequence for contig 13.
SEQ ID NO: 55 is the determined cDNA sequence for contig 14.
SEQ ID NO: 56 is the determined cDNA sequence for contig 15.
30 SEQ ID NO: 57 is the determined cDNA sequence for contig 16.
SEQ ID NO: 58 is the determined cDNA sequence for contig 17.

SEQ ID NO: 59 is the determined cDNA sequence for contig 18.
SEQ ID NO: 60 is the determined cDNA sequence for contig 19.
SEQ ID NO: 61 is the determined cDNA sequence for contig 20.
SEQ ID NO: 62 is the determined cDNA sequence for contig 21.
5 SEQ ID NO: 63 is the determined cDNA sequence for contig 22.
SEQ ID NO: 64 is the determined cDNA sequence for contig 23.
SEQ ID NO: 65 is the determined cDNA sequence for contig 24.
SEQ ID NO: 66 is the determined cDNA sequence for contig 25.
SEQ ID NO: 67 is the determined cDNA sequence for contig 26.
10 SEQ ID NO: 68 is the determined cDNA sequence for contig 27.
SEQ ID NO: 69 is the determined cDNA sequence for contig 28.
SEQ ID NO: 70 is the determined cDNA sequence for contig 29.
SEQ ID NO: 71 is the determined cDNA sequence for contig 30.
SEQ ID NO: 72 is the determined cDNA sequence for contig 31.
15 SEQ ID NO: 73 is the determined cDNA sequence for contig 32.
SEQ ID NO: 74 is the determined cDNA sequence for contig 33.
SEQ ID NO: 75 is the determined cDNA sequence for contig 34.
SEQ ID NO: 76 is the determined cDNA sequence for contig 35.
SEQ ID NO: 77 is the determined cDNA sequence for contig 36.
20 SEQ ID NO: 78 is the determined cDNA sequence for contig 37.
SEQ ID NO: 79 is the determined cDNA sequence for contig 38.
SEQ ID NO: 80 is the determined cDNA sequence for contig 39.
SEQ ID NO: 81 is the determined cDNA sequence for contig 40.
SEQ ID NO: 82 is the determined cDNA sequence for contig 41.
25 SEQ ID NO: 83 is the determined cDNA sequence for contig 42.
SEQ ID NO: 84 is the determined cDNA sequence for contig 43.
SEQ ID NO: 85 is the determined cDNA sequence for contig 44.
SEQ ID NO: 85 is the determined cDNA sequence for contig 45.
SEQ ID NO: 85 is the determined cDNA sequence for contig 46.
30 SEQ ID NO: 88 is the determined cDNA sequence for contig 47.
SEQ ID NO: 89 is the determined cDNA sequence for contig 48.

SEQ ID NO: 90 is the determined cDNA sequence for contig 49.
SEQ ID NO: 91 is the determined cDNA sequence for contig 50.
SEQ ID NO: 92 is the determined cDNA sequence for contig 51.
SEQ ID NO: 93 is the determined cDNA sequence for contig 52.
5 SEQ ID NO: 94 is the determined cDNA sequence for contig 53.
SEQ ID NO: 95 is the determined cDNA sequence for contig 54.
SEQ ID NO: 96 is the determined cDNA sequence for contig 55.
SEQ ID NO: 97 is the determined cDNA sequence for contig 56.
SEQ ID NO: 98 is the determined cDNA sequence for contig 57.
10 SEQ ID NO: 99 is the determined cDNA sequence for contig 58.
SEQ ID NO: 100 is the determined cDNA sequence for contig 59.
SEQ ID NO: 101 is the determined cDNA sequence for contig 60.
SEQ ID NO: 102 is the determined cDNA sequence for contig 61.
SEQ ID NO: 103 is the determined cDNA sequence for contig 62.
15 SEQ ID NO: 104 is the determined cDNA sequence for contig 63.
SEQ ID NO: 105 is the determined cDNA sequence for contig 64.
SEQ ID NO: 106 is the determined cDNA sequence for contig 65.
SEQ ID NO: 107 is the determined cDNA sequence for contig 66.
SEQ ID NO: 108 is the determined cDNA sequence for contig 67.
20 SEQ ID NO: 109 is the determined cDNA sequence for contig 68.
SEQ ID NO: 110 is the determined cDNA sequence for contig 69.
SEQ ID NO: 111 is the determined cDNA sequence for contig 70.
SEQ ID NO: 112 is the determined cDNA sequence for contig 71.
SEQ ID NO: 113 is the determined cDNA sequence for contig 72.
25 SEQ ID NO: 114 is the determined cDNA sequence for contig 73.
SEQ ID NO: 115 is the determined cDNA sequence for contig 74.
SEQ ID NO: 116 is the determined cDNA sequence for contig 75.
SEQ ID NO: 117 is the determined cDNA sequence for contig 76.
SEQ ID NO: 118 is the determined cDNA sequence for contig 77.
30 SEQ ID NO: 119 is the determined cDNA sequence for contig 78.
SEQ ID NO: 120 is the determined cDNA sequence for contig 79.

SEQ ID NO: 121 is the determined cDNA sequence for contig 80.
SEQ ID NO: 122 is the determined cDNA sequence for contig 81.
SEQ ID NO: 123 is the determined cDNA sequence for contig 82.
SEQ ID NO: 124 is the determined cDNA sequence for contig 83.
5 SEQ ID NO: 125 is the determined cDNA sequence for contig 84.
SEQ ID NO: 126 is the determined cDNA sequence for contig 85.
SEQ ID NO: 127 is the determined cDNA sequence for contig 86.
SEQ ID NO: 128 is the determined cDNA sequence for contig 87.
SEQ ID NO: 129 is the determined cDNA sequence for contig 88.
10 SEQ ID NO: 130 is the determined cDNA sequence for contig 89.
SEQ ID NO: 131 is the determined cDNA sequence for contig 90.
SEQ ID NO: 132 is the determined cDNA sequence for contig 91.
SEQ ID NO: 133 is the determined cDNA sequence for contig 92.
SEQ ID NO: 134 is the determined cDNA sequence for contig 93.
15 SEQ ID NO: 135 is the determined cDNA sequence for contig 94.
SEQ ID NO: 136 is the determined cDNA sequence for contig 95.
SEQ ID NO: 137 is the determined cDNA sequence for contig 96.
SEQ ID NO: 138 is the determined cDNA sequence for clone 47589.
SEQ ID NO: 139 is the determined cDNA sequence for clone 47578.
20 SEQ ID NO: 140 is the determined cDNA sequence for clone 47602.
SEQ ID NO: 141 is the determined cDNA sequence for clone 47593.
SEQ ID NO: 142 is the determined cDNA sequence for clone 47583.
SEQ ID NO: 143 is the determined cDNA sequence for clone 47624.
SEQ ID NO: 144 is the determined cDNA sequence for clone 47622.
25 SEQ ID NO: 145 is the determined cDNA sequence for clone 47649.
SEQ ID NO: 146 is the determined cDNA sequence for clone 48955.
SEQ ID NO: 147 is the determined cDNA sequence for clone 48962.
SEQ ID NO: 148 is the determined cDNA sequence for clone 48964.
SEQ ID NO: 149 is the determined cDNA sequence for clone 48987.
30 SEQ ID NO: 150 is the determined cDNA sequence for clone 49002.
SEQ ID NO: 151 is the determined cDNA sequence for clone 48950.

SEQ ID NO: 152 is the determined cDNA sequence for clone 48934.
SEQ ID NO: 153 is the determined cDNA sequence for clone 48960.
SEQ ID NO: 154 is the determined cDNA sequence for clone 48931.
SEQ ID NO: 155 is the determined cDNA sequence for clone 48935.
5 SEQ ID NO: 156 is the determined cDNA sequence for clone 48940.
SEQ ID NO: 157 is the determined cDNA sequence for clone 48936.
SEQ ID NO: 158 is the determined cDNA sequence for clone 48930.
SEQ ID NO: 159 is the determined cDNA sequence for clone 48956.
SEQ ID NO: 160 is the determined cDNA sequence for clone 48959.
10 SEQ ID NO: 161 is the determined cDNA sequence for clone 48949.
SEQ ID NO: 162 is the determined cDNA sequence for clone 48965.
SEQ ID NO: 163 is the determined cDNA sequence for clone 48970.
SEQ ID NO: 164 is the determined cDNA sequence for clone 48984.
SEQ ID NO: 165 is the determined cDNA sequence for clone 48969.
15 SEQ ID NO: 166 is the determined cDNA sequence for clone 48978.
SEQ ID NO: 167 is the determined cDNA sequence for clone 48968.
SEQ ID NO: 168 is the determined cDNA sequence for clone 48929.
SEQ ID NO: 169 is the determined cDNA sequence for clone 48937.
SEQ ID NO: 170 is the determined cDNA sequence for clone 48982.
20 SEQ ID NO: 171 is the determined cDNA sequence for clone 48983.
SEQ ID NO: 172 is the determined cDNA sequence for clone 48997.
SEQ ID NO: 173 is the determined cDNA sequence for clone 48992.
SEQ ID NO: 174 is the determined cDNA sequence for clone 49006.
SEQ ID NO: 175 is the determined cDNA sequence for clone 48994.
25 SEQ ID NO: 176 is the determined cDNA sequence for clone 49013.
SEQ ID NO: 177 is the determined cDNA sequence for clone 49008.
SEQ ID NO: 178 is the determined cDNA sequence for clone 48990.
SEQ ID NO: 179 is the determined cDNA sequence for clone 48989.
SEQ ID NO: 180 is the determined cDNA sequence for clone 49014.
30 SEQ ID NO: 181 is the determined cDNA sequence for clone 48988.
SEQ ID NO: 182 is the determined cDNA sequence for clone 49018.

SEQ ID NO: 183 is the determined cDNA sequence for clone 6921.
SEQ ID NO: 184 is the determined cDNA sequence for clone 6837.
SEQ ID NO: 185 is the determined cDNA sequence for clone 6840.
SEQ ID NO: 186 is the determined cDNA sequence for clone 6844.
5 SEQ ID NO: 187 is the determined cDNA sequence for clone 6854.
SEQ ID NO: 188 is the determined cDNA sequence for clone 6872.
SEQ ID NO: 189 is the determined cDNA sequence for clone 6906.
SEQ ID NO: 190 is the determined cDNA sequence for clone 6908.
SEQ ID NO: 191 is the determined cDNA sequence for clone 6910.
10 SEQ ID NO: 192 is the determined cDNA sequence for clone 6912.
SEQ ID NO: 193 is the determined cDNA sequence for clone 6913.
SEQ ID NO: 194 is the determined cDNA sequence for clone 6914.
SEQ ID NO: 195 is the determined cDNA sequence for clone 6916.
SEQ ID NO: 196 is the determined cDNA sequence for clone 6918.
15 SEQ ID NO: 197 is the determined cDNA sequence for clone 6924.
SEQ ID NO: 198 is the determined cDNA sequence for clone 6928.
SEQ ID NO: 199 is the determined cDNA sequence for clone 6978A.
SEQ ID NO: 200 is the determined cDNA sequence for clone 6978B.
SEQ ID NO: 201 is the determined cDNA sequence for clone 6982A.
20 SEQ ID NO: 202 is the determined cDNA sequence for clone 6982B.
SEQ ID NO: 203 is the determined cDNA sequence for clone 6850.
SEQ ID NO: 204 is the determined cDNA sequence for clone 6860.
SEQ ID NO: 205 is the determined cDNA sequence for O772P.
SEQ ID NO: 206 is the amino acid sequence encoded by SEQ ID NO:
25 205.
SEQ ID NO: 207 is the full-length cDNA sequence for O8E.
SEQ ID NO: 208 is a first amino acid sequence encoded by SEQ ID NO:
207.
SEQ ID NO: 209 is a second amino acid sequence encoded by SEQ ID
30 NO: 209.

SEQ ID NO: 210-290 are determined cDNA sequence of breast-tumor specific clones.

DETAILED DESCRIPTION OF THE INVENTION

As noted above, the present invention is generally directed to 5 compositions and methods for using the compositions, for example in the therapy and diagnosis of cancer, such as breast cancer. Certain illustrative compositions described herein include breast tumor polypeptides, polynucleotides encoding such polypeptides, binding agents such as antibodies, antigen presenting cells (APCs) and/or immune system cells (e.g., T cells). A "breast tumor protein," as the term is used herein, refers 10 generally to a protein that is expressed in breast tumor cells at a level that is at least two fold, and preferably at least five fold, greater than the level of expression in other normal tissues, as determined using a representative assay provided herein. Certain breast tumor proteins are tumor proteins that react detectably (within an immunoassay, such as an ELISA or Western blot) with antisera of a patient afflicted with breast 15 cancer.

Therefore, in accordance with the above, and as described further below, the present invention provides illustrative polynucleotide compositions having sequences set forth in SEQ ID NO:1-38, 42-204, 205, 207 and 210-290, polypeptides encoded by such polynucleotides, antibody compositions capable of binding such 20 polypeptides, and numerous additional embodiments employing such compositions, for example in the detection, diagnosis and/or therapy of human breast cancer.

POLYNUCLEOTIDE COMPOSITIONS

As used herein, the terms "DNA segment" and "polynucleotide" refer to a DNA molecule that has been isolated free of total genomic DNA of a particular 25 species. Therefore, a DNA segment encoding a polypeptide refers to a DNA segment that contains one or more coding sequences yet is substantially isolated away from, or purified free from, total genomic DNA of the species from which the DNA segment is obtained. Included within the terms "DNA segment" and "polynucleotide" are DNA

segments and smaller fragments of such segments, and also recombinant vectors, including, for example, plasmids, cosmids, phagemids, phage, viruses, and the like.

As will be understood by those skilled in the art, the DNA segments of this invention can include genomic sequences, extra-genomic and plasmid-encoded 5 sequences and smaller engineered gene segments that express, or may be adapted to express, proteins, polypeptides, peptides and the like. Such segments may be naturally isolated, or modified synthetically by the hand of man.

"Isolated," as used herein, means that a polynucleotide is substantially away from other coding sequences, and that the DNA segment does not contain large 10 portions of unrelated coding DNA, such as large chromosomal fragments or other functional genes or polypeptide coding regions. Of course, this refers to the DNA segment as originally isolated, and does not exclude genes or coding regions later added to the segment by the hand of man.

As will be recognized by the skilled artisan, polynucleotides may be 15 single-stranded (coding or antisense) or double-stranded, and may be DNA (genomic, cDNA or synthetic) or RNA molecules. RNA molecules include HnRNA molecules, which contain introns and correspond to a DNA molecule in a one-to-one manner, and mRNA molecules, which do not contain introns. Additional coding or non-coding sequences may, but need not, be present within a polynucleotide of the present 20 invention, and a polynucleotide may, but need not, be linked to other molecules and/or support materials.

Polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence that encodes a breast tumor protein or a portion thereof) or may comprise a variant, or a biological or antigenic functional equivalent of such a sequence. 25 Polynucleotide variants may contain one or more substitutions, additions, deletions and/or insertions, as further described below, preferably such that the immunogenicity of the encoded polypeptide is not diminished, relative to a native tumor protein. The effect on the immunogenicity of the encoded polypeptide may generally be assessed as described herein. The term "variants" also encompasses homologous genes of 30 xenogenic origin.

When comparing polynucleotide or polypeptide sequences, two sequences are said to be "identical" if the sequence of nucleotides or amino acids in the two sequences is the same when aligned for maximum correspondence, as described below. Comparisons between two sequences are typically performed by comparing the 5 sequences over a comparison window to identify and compare local regions of sequence similarity. A "comparison window" as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, 40 to about 50, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

10 Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies several alignment schemes described in the following references: Dayhoff, M.O. (1978) A model of evolutionary change in proteins – Matrices for detecting distant relationships.

15 In Dayhoff, M.O. (ed.) *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, Washington DC Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990) *Unified Approach to Alignment and Phylogenies* pp. 626-645 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989) *CABIOS* 5:151-153; Myers, E.W. and Muller W. (1988) *CABIOS* 4:11-17; Robinson, 20 E.D. (1971) *Comb. Theor* 11:105; Santou, N. Nes, M. (1987) *Mol. Biol. Evol.* 4:406-425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy – the Principles and Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and Lipman, D.J. (1983) *Proc. Natl. Acad. Sci. USA* 80:726-730.

25 Alternatively, optimal alignment of sequences for comparison may be conducted by the local identity algorithm of Smith and Waterman (1981) *Add. APL. Math* 2:482, by the identity alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443, by the search for similarity methods of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics 30 Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

One preferred example of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al. (1977) *Nucl. Acids Res.* 25:3389-3402 and Altschul et al. (1990) *J. Mol. Biol.* 215:403-410, respectively. BLAST and BLAST 2.0 can be used, for example with the parameters described herein, to determine percent sequence identity for the polynucleotides and polypeptides of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. In one illustrative example, cumulative scores can be calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always <0). For amino acid sequences, a scoring matrix can be used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915) alignments, (B) of 50, expectation (E) of 10, M=5, N=-4 and a comparison of both strands.

Preferably, the "percentage of sequence identity" is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polynucleotide or polypeptide sequence in the comparison window may comprise additions or deletions (i.e., gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid bases or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the reference sequence (i.e., the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

Therefore, the present invention encompasses polynucleotide and polypeptide sequences having substantial identity to the sequences disclosed herein, for example those comprising at least 50% sequence identity, preferably at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% or higher, sequence 5 identity compared to a polynucleotide or polypeptide sequence of this invention using the methods described herein, (e.g., BLAST analysis using standard parameters, as described below). One skilled in this art will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, 10 reading frame positioning and the like.

In additional embodiments, the present invention provides isolated polynucleotides and polypeptides comprising various lengths of contiguous stretches of sequence identical to or complementary to one or more of the sequences disclosed herein. For example, polynucleotides are provided by this invention that comprise at 15 least about 15, 20, 30, 40, 50, 75, 100, 150, 200, 300, 400, 500 or 1000 or more contiguous nucleotides of one or more of the sequences disclosed herein as well as all intermediate lengths there between. It will be readily understood that "intermediate lengths", in this context, means any length between the quoted values, such as 16, 17, 18, 19, *etc.*; 21, 22, 23, *etc.*; 30, 31, 32, *etc.*; 50, 51, 52, 53, *etc.*; 100, 101, 102, 103, 20 *etc.*; 150, 151, 152, 153, *etc.*; including all integers through 200-500; 500-1,000, and the like.

The polynucleotides of the present invention, or fragments thereof, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters, polyadenylation signals, additional restriction 25 enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length may vary considerably. It is therefore contemplated that a nucleic acid fragment of almost any length may be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol. For example, illustrative DNA segments with total lengths of about 10,000, about 5000, 30 about 3000, about 2,000, about 1,000, about 500, about 200, about 100, about 50 base

pairs in length, and the like, (including all intermediate lengths) are contemplated to be useful in many implementations of this invention.

In other embodiments, the present invention is directed to polynucleotides that are capable of hybridizing under moderately stringent conditions to 5 a polynucleotide sequence provided herein, or a fragment thereof, or a complementary sequence thereof. Hybridization techniques are well known in the art of molecular biology. For purposes of illustration, suitable moderately stringent conditions for testing the hybridization of a polynucleotide of this invention with other polynucleotides include prewashing in a solution of 5 X SSC, 0.5% SDS, 1.0 mM 10 EDTA (pH 8.0); hybridizing at 50°C-65°C, 5 X SSC, overnight; followed by washing twice at 65°C for 20 minutes with each of 2X, 0.5X and 0.2X SSC containing 0.1% SDS.

Moreover, it will be appreciated by those of ordinary skill in the art that, as a result of the degeneracy of the genetic code, there are many nucleotide sequences 15 that encode a polypeptide as described herein. Some of these polynucleotides bear minimal homology to the nucleotide sequence of any native gene. Nonetheless, polynucleotides that vary due to differences in codon usage are specifically contemplated by the present invention. Further, alleles of the genes comprising the polynucleotide sequences provided herein are within the scope of the present invention. 20 Alleles are endogenous genes that are altered as a result of one or more mutations, such as deletions, additions and/or substitutions of nucleotides. The resulting mRNA and protein may, but need not, have an altered structure or function. Alleles may be identified using standard techniques (such as hybridization, amplification and/or database sequence comparison).

25 PROBES AND PRIMERS

In other embodiments of the present invention, the polynucleotide sequences provided herein can be advantageously used as probes or primers for nucleic acid hybridization. As such, it is contemplated that nucleic acid segments that comprise a sequence region of at least about 15 nucleotide long contiguous sequence that has the 30 same sequence as, or is complementary to, a 15 nucleotide long contiguous sequence

disclosed herein will find particular utility. Longer contiguous identical or complementary sequences, *e.g.*, those of about 20, 30, 40, 50, 100, 200, 500, 1000 (including all intermediate lengths) and even up to full length sequences will also be of use in certain embodiments.

5 The ability of such nucleic acid probes to specifically hybridize to a sequence of interest will enable them to be of use in detecting the presence of complementary sequences in a given sample. However, other uses are also envisioned, such as the use of the sequence information for the preparation of mutant species primers, or primers for use in preparing other genetic constructions.

10 Polynucleotide molecules having sequence regions consisting of contiguous nucleotide stretches of 10-14, 15-20, 30, 50, or even of 100-200 nucleotides or so (including intermediate lengths as well), identical or complementary to a polynucleotide sequence disclosed herein, are particularly contemplated as hybridization probes for use in, *e.g.*, Southern and Northern blotting. This would allow
15 a gene product, or fragment thereof, to be analyzed, both in diverse cell types and also in various bacterial cells. The total size of fragment, as well as the size of the complementary stretch(es), will ultimately depend on the intended use or application of the particular nucleic acid segment. Smaller fragments will generally find use in hybridization embodiments, wherein the length of the contiguous complementary
20 region may be varied, such as between about 15 and about 100 nucleotides, but larger contiguous complementarity stretches may be used, according to the length complementary sequences one wishes to detect.

25 The use of a hybridization probe of about 15-25 nucleotides in length allows the formation of a duplex molecule that is both stable and selective. Molecules having contiguous complementary sequences over stretches greater than 15 bases in length are generally preferred, though, in order to increase stability and selectivity of the hybrid, and thereby improve the quality and degree of specific hybrid molecules obtained. One will generally prefer to design nucleic acid molecules having gene-complementary stretches of 15 to 25 contiguous nucleotides, or even longer where
30 desired.

Hybridization probes may be selected from any portion of any of the sequences disclosed herein. All that is required is to review the sequence set forth in SEQ ID NO: 1-38, 42-204, 205, 207 and 210-290, or to any continuous portion of the sequence, from about 15-25 nucleotides in length up to and including the full length 5 sequence, that one wishes to utilize as a probe or primer. The choice of probe and primer sequences may be governed by various factors. For example, one may wish to employ primers from towards the termini of the total sequence.

Small polynucleotide segments or fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means, as is commonly 10 practiced using an automated oligonucleotide synthesizer. Also, fragments may be obtained by application of nucleic acid reproduction technology, such as the PCR™ technology of U. S. Patent 4,683,202 (incorporated herein by reference), by introducing selected sequences into recombinant vectors for recombinant production, and by other recombinant DNA techniques generally known to those of skill in the art of molecular 15 biology.

The nucleotide sequences of the invention may be used for their ability to selectively form duplex molecules with complementary stretches of the entire gene or gene fragments of interest. Depending on the application envisioned, one will typically desire to employ varying conditions of hybridization to achieve varying degrees of 20 selectivity of probe towards target sequence. For applications requiring high selectivity, one will typically desire to employ relatively stringent conditions to form the hybrids, e.g., one will select relatively low salt and/or high temperature conditions, such as provided by a salt concentration of from about 0.02 M to about 0.15 M salt at temperatures of from about 50°C to about 70°C. Such selective conditions tolerate 25 little, if any, mismatch between the probe and the template or target strand, and would be particularly suitable for isolating related sequences.

Of course, for some applications, for example, where one desires to prepare mutants employing a mutant primer strand hybridized to an underlying template, less stringent (reduced stringency) hybridization conditions will typically be 30 needed in order to allow formation of the heteroduplex. In these circumstances, one may desire to employ salt conditions such as those of from about 0.15 M to about 0.9 M

salt, at temperatures ranging from about 20°C to about 55°C. Cross-hybridizing species can thereby be readily identified as positively hybridizing signals with respect to control hybridizations. In any case, it is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide, which serves to 5 destabilize the hybrid duplex in the same manner as increased temperature. Thus, hybridization conditions can be readily manipulated, and thus will generally be a method of choice depending on the desired results.

POLYNUCLEOTIDE IDENTIFICATION AND CHARACTERIZATION

Polynucleotides may be identified, prepared and/or manipulated using 10 any of a variety of well established techniques. For example, a polynucleotide may be identified, as described in more detail below, by screening a microarray of cDNAs for tumor-associated expression (*i.e.*, expression that is at least two fold greater in a tumor than in normal tissue, as determined using a representative assay provided herein). Such screens may be performed, for example, using a Synteni microarray (Palo Alto, 15 CA) according to the manufacturer's instructions (and essentially as described by Schena et al., *Proc. Natl. Acad. Sci. USA* 93:10614-10619, 1996 and Heller et al., *Proc. Natl. Acad. Sci. USA* 94:2150-2155, 1997). Alternatively, polynucleotides may be amplified from cDNA prepared from cells expressing the proteins described herein, such as breast tumor cells. Such polynucleotides may be amplified via polymerase 20 chain reaction (PCR). For this approach, sequence-specific primers may be designed based on the sequences provided herein, and may be purchased or synthesized.

An amplified portion of a polynucleotide of the present invention may be used to isolate a full length gene from a suitable library (*e.g.*, a breast tumor cDNA library) using well known techniques. Within such techniques, a library (cDNA or 25 genomic) is screened using one or more polynucleotide probes or primers suitable for amplification. Preferably, a library is size-selected to include larger molecules. Random primed libraries may also be preferred for identifying 5' and upstream regions of genes. Genomic libraries are preferred for obtaining introns and extending 5' sequences.

For hybridization techniques, a partial sequence may be labeled (e.g., by nick-translation or end-labeling with ^{32}P) using well known techniques. A bacterial or bacteriophage library is then generally screened by hybridizing filters containing denatured bacterial colonies (or lawns containing phage plaques) with the labeled probe

5 (see Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989). Hybridizing colonies or plaques are selected and expanded, and the DNA is isolated for further analysis. cDNA clones may be analyzed to determine the amount of additional sequence by, for example, PCR using a primer from the partial sequence and a primer from the vector. Restriction maps and

10 partial sequences may be generated to identify one or more overlapping clones. The complete sequence may then be determined using standard techniques, which may involve generating a series of deletion clones. The resulting overlapping sequences can then be assembled into a single contiguous sequence. A full length cDNA molecule can be generated by ligating suitable fragments, using well known techniques.

15 Alternatively, there are numerous amplification techniques for obtaining a full length coding sequence from a partial cDNA sequence. Within such techniques, amplification is generally performed via PCR. Any of a variety of commercially available kits may be used to perform the amplification step. Primers may be designed using, for example, software well known in the art. Primers are preferably 22-30

20 nucleotides in length, have a GC content of at least 50% and anneal to the target sequence at temperatures of about 68°C to 72°C. The amplified region may be sequenced as described above, and overlapping sequences assembled into a contiguous sequence.

One such amplification technique is inverse PCR (see Triglia et al., *Nucl. Acids Res.* 16:8186, 1988), which uses restriction enzymes to generate a fragment in the known region of the gene. The fragment is then circularized by intramolecular ligation and used as a template for PCR with divergent primers derived from the known region. Within an alternative approach, sequences adjacent to a partial sequence may be retrieved by amplification with a primer to a linker sequence and a primer specific to a

30 known region. The amplified sequences are typically subjected to a second round of amplification with the same linker primer and a second primer specific to the known

region. A variation on this procedure, which employs two primers that initiate extension in opposite directions from the known sequence, is described in WO 96/38591. Another such technique is known as "rapid amplification of cDNA ends" or RACE. This technique involves the use of an internal primer and an external primer,
5 which hybridizes to a polyA region or vector sequence, to identify sequences that are 5' and 3' of a known sequence. Additional techniques include capture PCR (Lagerstrom et al., *PCR Methods Applic.* 1:111-19, 1991) and walking PCR (Parker et al., *Nucl. Acids. Res.* 19:3055-60, 1991). Other methods employing amplification may also be employed to obtain a full length cDNA sequence.

10 In certain instances, it is possible to obtain a full length cDNA sequence by analysis of sequences provided in an expressed sequence tag (EST) database, such as that available from GenBank. Searches for overlapping ESTs may generally be performed using well known programs (e.g., NCBI BLAST searches), and such ESTs may be used to generate a contiguous full length sequence. Full length DNA sequences
15 may also be obtained by analysis of genomic fragments.

POLYNUCLEOTIDE EXPRESSION IN HOST CELLS

In other embodiments of the invention, polynucleotide sequences or fragments thereof which encode polypeptides of the invention, or fusion proteins or functional equivalents thereof, may be used in recombinant DNA molecules to direct
20 expression of a polypeptide in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences that encode substantially the same or a functionally equivalent amino acid sequence may be produced and these sequences may be used to clone and express a given polypeptide.

As will be understood by those of skill in the art, it may be advantageous
25 in some instances to produce polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce a recombinant RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring
30 sequence.

Moreover, the polynucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter polypeptide encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the gene product. For 5 example, DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. In addition, site-directed mutagenesis may be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, or introduce mutations, and so forth.

10 In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences may be ligated to a heterologous sequence to encode a fusion protein. For example, to screen peptide libraries for inhibitors of polypeptide activity, it may be useful to encode a chimeric protein that can be recognized by a commercially available antibody. A fusion protein may also be 15 engineered to contain a cleavage site located between the polypeptide-encoding sequence and the heterologous protein sequence, so that the polypeptide may be cleaved and purified away from the heterologous moiety.

Sequences encoding a desired polypeptide may be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers, M. H. et al. 20 (1980) *Nucl. Acids Res. Symp. Ser.* 215-223, Horn, T. et al. (1980) *Nucl. Acids Res. Symp. Ser.* 225-232). Alternatively, the protein itself may be produced using chemical methods to synthesize the amino acid sequence of a polypeptide, or a portion thereof. For example, peptide synthesis can be performed using various solid-phase techniques (Roberge, J. Y. et al. (1995) *Science* 269:202-204) and automated synthesis may be 25 achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer, Palo Alto, CA).

A newly synthesized peptide may be substantially purified by preparative high performance liquid chromatography (e.g., Creighton, T. (1983) Proteins, Structures and Molecular Principles, WH Freeman and Co., New York, N.Y.) 30 or other comparable techniques available in the art. The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (e.g., the Edman

degradation procedure). Additionally, the amino acid sequence of a polypeptide, or any part thereof, may be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

5 In order to express a desired polypeptide, the nucleotide sequences encoding the polypeptide, or functional equivalents, may be inserted into appropriate expression vector, i.e., a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well known to those skilled in the art may be used to construct expression vectors containing
10 sequences encoding a polypeptide of interest and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described in Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, N.Y., and Ausubel, F. M. et al. (1989) Current
15 Protocols in Molecular Biology, John Wiley & Sons, New York. N.Y.

A variety of expression vector/host systems may be utilized to contain and express polynucleotide sequences. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; 20 insect cell systems infected with virus expression vectors (e.g., baculovirus); plant cell systems transformed with virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems.

25 The "control elements" or "regulatory sequences" present in an expression vector are those non-translated regions of the vector--enhancers, promoters, 5' and 3' untranslated regions--which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used.
30 For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the PBLUESCRIPT phagemid (Stratagene, La Jolla, Calif.) or

PSPORT1 plasmid (Gibco BRL, Gaithersburg, MD) and the like may be used. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are generally preferred. If it is necessary to generate a cell line that contains multiple copies of the sequence encoding a polypeptide, vectors based on SV40 or EBV 5 may be advantageously used with an appropriate selectable marker.

In bacterial systems, a number of expression vectors may be selected depending upon the use intended for the expressed polypeptide. For example, when large quantities are needed, for example for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be used: 10 Such vectors include, but are not limited to, the multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene), in which the sequence encoding the polypeptide of interest may be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of .beta.-galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke, G. and S. M. Schuster (1989) *J. 15 Biol. Chem.* 264:5503-5509); and the like. pGEX Vectors (Promega, Madison, Wis.) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems may be designed to include 20 heparin, thrombin, or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast, *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. For reviews, see Ausubel et al. (supra) and Grant et al. (1987) *Methods 25 Enzymol.* 153:516-544.

In cases where plant expression vectors are used, the expression of sequences encoding polypeptides may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV may be used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. 30 (1987) *EMBO J.* 6:307-311. Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. et al. (1984) *EMBO J.*

3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (see, for example, Hobbs, S. or 5 Murry, L. E. in McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York, N.Y.; pp. 191-196).

An insect system may also be used to express a polypeptide of interest. For example, in one such system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or 10 in *Trichoplusia* larvae. The sequences encoding the polypeptide may be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of the polypeptide-encoding sequence will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses may then be used to infect, for example, *S. frugiperda* 15 cells or *Trichoplusia* larvae in which the polypeptide of interest may be expressed (Engelhard, E. K. et al. (1994) *Proc. Natl. Acad. Sci.* 91:3224-3227).

In mammalian host cells, a number of viral-based expression systems are generally available. For example, in cases where an adenovirus is used as an expression vector, sequences encoding a polypeptide of interest may be ligated into an adenovirus 20 transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain a viable virus which is capable of expressing the polypeptide in infected host cells (Logan, J. and Shenk, T. (1984) *Proc. Natl. Acad. Sci.* 81:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used 25 to increase expression in mammalian host cells.

Specific initiation signals may also be used to achieve more efficient translation of sequences encoding a polypeptide of interest. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding the polypeptide, its initiation codon, and upstream sequences are inserted into the 30 appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a portion

thereof, is inserted, exogenous translational control signals including the ATG initiation codon should be provided. Furthermore, the initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic.

5 The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used, such as those described in the literature (Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162).

In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the 10 desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells such as CHO, HeLa, MDCK, HEK293, and WI38, which have specific cellular machinery and 15 characteristic mechanisms for such post-translational activities, may be chosen to ensure the correct modification and processing of the foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is generally preferred. For example, cell lines which stably express a polynucleotide of interest may be transformed using expression vectors which may 20 contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which 25 successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler, M. et al. (1977) *Cell* 11:223-32) and adenine phosphoribosyltransferase 30 (Lowy, I. et al. (1990) *Cell* 22:817-23) genes which can be employed in tk.sup.- or aprt.sup.- cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can

be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler, M. et al. (1980) *Proc. Natl. Acad. Sci.* 77:3567-70); npt, which confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin, F. et al (1981) *J. Mol. Biol.* 150:1-14); and als or pat, which confer resistance to 5 chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, *supra*). Additional selectable genes have been described, for example, trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman, S. C. and R. C. Mulligan (1988) *Proc. Natl. Acad. Sci.* 85:8047-51). Recently, the use of visible markers has gained popularity with such 10 markers as anthocyanins, beta-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C. A. et al. (1995) *Methods Mol. Biol.* 55:121-131).

Although the presence/absence of marker gene expression suggests that 15 the gene of interest is also present, its presence and expression may need to be confirmed. For example, if the sequence encoding a polypeptide is inserted within a marker gene sequence, recombinant cells containing sequences can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a polypeptide-encoding sequence under the control of a single promoter. 20 Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

Alternatively, host cells which contain and express a desired polynucleotide sequence may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA- 25 RNA hybridizations and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein.

A variety of protocols for detecting and measuring the expression of 30 polynucleotide-encoded products, using either polyclonal or monoclonal antibodies specific for the product are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated

cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on a given polypeptide may be preferred for some applications, but a competitive binding assay may also be employed. These and other assays are described, among other places, in Hampton, R. et al. (1990; 5 Serological Methods, a Laboratory Manual, APS Press, St Paul, Minn.) and Maddox, D. E. et al. (1983; *J. Exp. Med.* 158:1211-1216).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to 10 polynucleotides include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences, or any portions thereof may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 15 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits. Suitable reporter molecules or labels, which may be used include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with a polynucleotide sequence of interest may be 20 cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides of the invention may be designed to contain signal sequences which direct secretion of the 25 encoded polypeptide through a prokaryotic or eukaryotic cell membrane. Other recombinant constructions may be used to join sequences encoding a polypeptide of interest to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow 30 purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAG extension/affinity

purification system (Immunex Corp., Seattle, Wash.). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen, San Diego, Calif.) between the purification domain and the encoded polypeptide may be used to facilitate purification. One such expression vector provides for expression of a fusion 5 protein containing a polypeptide of interest and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC (immobilized metal ion affinity chromatography) as described in Porath, J. et al. (1992, *Prot. Exp. Purif.* 3:263-281) while the enterokinase cleavage site provides a means for purifying the desired polypeptide from the fusion 10 protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D. J. et al. (1993; *DNA Cell Biol.* 12:441-453).

In addition to recombinant production methods, polypeptides of the invention, and fragments thereof, may be produced by direct peptide synthesis using solid-phase techniques (Merrifield J. (1963) *J. Am. Chem. Soc.* 85:2149-2154). Protein 15 synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Alternatively, various fragments may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

20 SITE-SPECIFIC MUTAGENESIS

Site-specific mutagenesis is a technique useful in the preparation of individual peptides, or biologically functional equivalent polypeptides, through specific mutagenesis of the underlying polynucleotides that encode them. The technique, well-known to those of skill in the art, further provides a ready ability to prepare and 25 test sequence variants, for example, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the DNA. Site-specific mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of 30 sufficient size and sequence complexity to form a stable duplex on both sides of the

deletion junction being traversed. Mutations may be employed in a selected polynucleotide sequence to improve, alter, decrease, modify, or otherwise change the properties of the polynucleotide itself, and/or alter the properties, activity, composition, stability, or primary sequence of the encoded polypeptide.

5 In certain embodiments of the present invention, the inventors contemplate the mutagenesis of the disclosed polynucleotide sequences to alter one or more properties of the encoded polypeptide, such as the antigenicity of a polypeptide vaccine. The techniques of site-specific mutagenesis are well-known in the art, and are widely used to create variants of both polypeptides and polynucleotides. For example, 10 site-specific mutagenesis is often used to alter a specific portion of a DNA molecule. In such embodiments, a primer comprising typically about 14 to about 25 nucleotides or so in length is employed, with about 5 to about 10 residues on both sides of the junction of the sequence being altered.

15 As will be appreciated by those of skill in the art, site-specific mutagenesis techniques have often employed a phage vector that exists in both a single stranded and double stranded form. Typical vectors useful in site-directed mutagenesis include vectors such as the M13 phage. These phage are readily commercially-available and their use is generally well-known to those skilled in the art. Double-stranded plasmids are also routinely employed in site directed mutagenesis that 20 eliminates the step of transferring the gene of interest from a plasmid to a phage.

25 In general, site-directed mutagenesis in accordance herewith is performed by first obtaining a single-stranded vector or melting apart of two strands of a double-stranded vector that includes within its sequence a DNA sequence that encodes the desired peptide. An oligonucleotide primer bearing the desired mutated sequence is prepared, generally synthetically. This primer is then annealed with the single-stranded vector, and subjected to DNA polymerizing enzymes such as *E. coli* polymerase I Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the original non-mutated sequence and the second strand bears the desired mutation. This heteroduplex vector is 30 then used to transform appropriate cells, such as *E. coli* cells, and clones are selected which include recombinant vectors bearing the mutated sequence arrangement.

The preparation of sequence variants of the selected peptide-encoding DNA segments using site-directed mutagenesis provides a means of producing potentially useful species and is not meant to be limiting as there are other ways in which sequence variants of peptides and the DNA sequences encoding them may be obtained. For example, recombinant vectors encoding the desired peptide sequence may be treated with mutagenic agents, such as hydroxylamine, to obtain sequence variants. Specific details regarding these methods and protocols are found in the teachings of Maloy *et al.*, 1994; Segal, 1976; Prokop and Bajpai, 1991; Kuby, 1994; and Maniatis *et al.*, 1982, each incorporated herein by reference, for that purpose.

As used herein, the term "oligonucleotide directed mutagenesis procedure" refers to template-dependent processes and vector-mediated propagation which result in an increase in the concentration of a specific nucleic acid molecule relative to its initial concentration, or in an increase in the concentration of a detectable signal, such as amplification. As used herein, the term "oligonucleotide directed mutagenesis procedure" is intended to refer to a process that involves the template-dependent extension of a primer molecule. The term template dependent process refers to nucleic acid synthesis of an RNA or a DNA molecule wherein the sequence of the newly synthesized strand of nucleic acid is dictated by the well-known rules of complementary base pairing (see, for example, Watson, 1987). Typically, vector mediated methodologies involve the introduction of the nucleic acid fragment into a DNA or RNA vector, the clonal amplification of the vector, and the recovery of the amplified nucleic acid fragment. Examples of such methodologies are provided by U. S. Patent No. 4,237,224, specifically incorporated herein by reference in its entirety.

POLYNUCLEOTIDE AMPLIFICATION TECHNIQUES

A number of template dependent processes are available to amplify the target sequences of interest present in a sample. One of the best known amplification methods is the polymerase chain reaction (PCRTM) which is described in detail in U.S. Patent Nos. 4,683,195, 4,683,202 and 4,800,159, each of which is incorporated herein by reference in its entirety. Briefly, in PCRTM, two primer sequences are prepared which are complementary to regions on opposite complementary strands of the target

sequence. An excess of deoxynucleoside triphosphates is added to a reaction mixture along with a DNA polymerase (e.g., *Taq* polymerase). If the target sequence is present in a sample, the primers will bind to the target and the polymerase will cause the primers to be extended along the target sequence by adding on nucleotides. By raising 5 and lowering the temperature of the reaction mixture, the extended primers will dissociate from the target to form reaction products, excess primers will bind to the target and to the reaction product and the process is repeated. Preferably reverse transcription and PCR™ amplification procedure may be performed in order to quantify the amount of mRNA amplified. Polymerase chain reaction methodologies are well 10 known in the art.

Another method for amplification is the ligase chain reaction (referred to as LCR), disclosed in Eur. Pat. Appl. Publ. No. 320,308 (specifically incorporated herein by reference in its entirety). In LCR, two complementary probe pairs are prepared, and in the presence of the target sequence, each pair will bind to opposite 15 complementary strands of the target such that they abut. In the presence of a ligase, the two probe pairs will link to form a single unit. By temperature cycling, as in PCR™, bound ligated units dissociate from the target and then serve as "target sequences" for ligation of excess probe pairs. U.S. Patent No. 4,883,750, incorporated herein by reference in its entirety, describes an alternative method of amplification similar to LCR 20 for binding probe pairs to a target sequence.

Qbeta Replicase, described in PCT Intl. Pat. Appl. Publ. No. PCT/US87/00880, incorporated herein by reference in its entirety, may also be used as still another amplification method in the present invention. In this method, a replicative sequence of RNA that has a region complementary to that of a target is added to a 25 sample in the presence of an RNA polymerase. The polymerase will copy the replicative sequence that can then be detected.

An isothermal amplification method, in which restriction endonucleases and ligases are used to achieve the amplification of target molecules that contain nucleotide 5'-[α -thio]triphosphates in one strand of a restriction site (Walker *et al.*, 30 1992, incorporated herein by reference in its entirety), may also be useful in the amplification of nucleic acids in the present invention.

Strand Displacement Amplification (SDA) is another method of carrying out isothermal amplification of nucleic acids which involves multiple rounds of strand displacement and synthesis, *i.e.* nick translation. A similar method, called Repair Chain Reaction (RCR) is another method of amplification which may be useful in the present 5 invention and is involves annealing several probes throughout a region targeted for amplification, followed by a repair reaction in which only two of the four bases are present. The other two bases can be added as biotinylated derivatives for easy detection. A similar approach is used in SDA.

Sequences can also be detected using a cyclic probe reaction (CPR). In 10 CPR, a probe having a 3' and 5' sequences of non-target DNA and an internal or "middle" sequence of the target protein specific RNA is hybridized to DNA which is present in a sample. Upon hybridization, the reaction is treated with RNaseH, and the products of the probe are identified as distinctive products by generating a signal that is released after digestion. The original template is annealed to another cycling probe and 15 the reaction is repeated. Thus, CPR involves amplifying a signal generated by hybridization of a probe to a target gene specific expressed nucleic acid.

Still other amplification methods described in Great Britain Pat. Appl. No. 2 202 328, and in PCT Intl. Pat. Appl. Publ. No. PCT/US89/01025, each of which is incorporated herein by reference in its entirety, may be used in accordance with the 20 present invention. In the former application, "modified" primers are used in a PCR-like, template and enzyme dependent synthesis. The primers may be modified by labeling with a capture moiety (*e.g.*, biotin) and/or a detector moiety (*e.g.*, enzyme). In the latter application, an excess of labeled probes is added to a sample. In the presence of the target sequence, the probe binds and is cleaved catalytically. After cleavage, the target 25 sequence is released intact to be bound by excess probe. Cleavage of the labeled probe signals the presence of the target sequence.

Other nucleic acid amplification procedures include transcription-based amplification systems (TAS) (Kwoh *et al.*, 1989; PCT Intl. Pat. Appl. Publ. No. WO 88/10315, incorporated herein by reference in its entirety), including nucleic acid 30 sequence based amplification (NASBA) and 3SR. In NASBA, the nucleic acids can be prepared for amplification by standard phenol/chloroform extraction, heat denaturation

of a sample, treatment with lysis buffer and minispin columns for isolation of DNA and RNA or guanidinium chloride extraction of RNA. These amplification techniques involve annealing a primer that has sequences specific to the target sequence. Following polymerization, DNA/RNA hybrids are digested with RNase H while double 5 stranded DNA molecules are heat-denatured again. In either case the single stranded DNA is made fully double stranded by addition of second target-specific primer, followed by polymerization. The double stranded DNA molecules are then multiply transcribed by a polymerase such as T7 or SP6. In an isothermal cyclic reaction, the RNAs are reverse transcribed into DNA, and transcribed once again with a polymerase 10 such as T7 or SP6. The resulting products, whether truncated or complete, indicate target-specific sequences.

Eur. Pat. Appl. Publ. No. 329,822, incorporated herein by reference in its entirety, disclose a nucleic acid amplification process involving cyclically synthesizing single-stranded RNA ("ssRNA"), ssDNA, and double-stranded DNA (dsDNA), which 15 may be used in accordance with the present invention. The ssRNA is a first template for a first primer oligonucleotide, which is elongated by reverse transcriptase (RNA-dependent DNA polymerase). The RNA is then removed from resulting DNA:RNA duplex by the action of ribonuclease H (RNase H, an RNase specific for RNA in a duplex with either DNA or RNA). The resultant ssDNA is a second template 20 for a second primer, which also includes the sequences of an RNA polymerase promoter (exemplified by T7 RNA polymerase) 5' to its homology to its template. This primer is then extended by DNA polymerase (exemplified by the large "Klenow" fragment of *E. coli* DNA polymerase I), resulting as a double-stranded DNA ("dsDNA") molecule, having a sequence identical to that of the original RNA between the primers and having 25 additionally, at one end, a promoter sequence. This promoter sequence can be used by the appropriate RNA polymerase to make many RNA copies of the DNA. These copies can then re-enter the cycle leading to very swift amplification. With proper choice of enzymes, this amplification can be done isothermally without addition of enzymes at each cycle. Because of the cyclical nature of this process, the starting sequence can be 30 chosen to be in the form of either DNA or RNA.

PCT Intl. Pat. Appl. Publ. No. WO 89/06700, incorporated herein by reference in its entirety, disclose a nucleic acid sequence amplification scheme based on the hybridization of a promoter/primer sequence to a target single-stranded DNA ("ssDNA") followed by transcription of many RNA copies of the sequence. This 5 scheme is not cyclic; i.e. new templates are not produced from the resultant RNA transcripts. Other amplification methods include "RACE" (Frohman, 1990), and "one-sided PCR" (Ohara, 1989) which are well-known to those of skill in the art.

Methods based on ligation of two (or more) oligonucleotides in the presence of nucleic acid having the sequence of the resulting "di-oligonucleotide", 10 thereby amplifying the di-oligonucleotide (Wu and Dean, 1996, incorporated herein by reference in its entirety), may also be used in the amplification of DNA sequences of the present invention.

BIOLOGICAL FUNCTIONAL EQUIVALENTS

Modification and changes may be made in the structure of the 15 polynucleotides and polypeptides of the present invention and still obtain a functional molecule that encodes a polypeptide with desirable characteristics. As mentioned above, it is often desirable to introduce one or more mutations into a specific polynucleotide sequence. In certain circumstances, the resulting encoded polypeptide sequence is altered by this mutation, or in other cases, the sequence of the polypeptide 20 is unchanged by one or more mutations in the encoding polynucleotide.

When it is desirable to alter the amino acid sequence of a polypeptide to create an equivalent, or even an improved, second-generation molecule, the amino acid changes may be achieved by changing one or more of the codons of the encoding DNA sequence, according to Table 1.

25 For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate molecules. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence 30 substitutions can be made in a protein sequence, and, of course, its underlying DNA

coding sequence, and nevertheless obtain a protein with like properties. It is thus contemplated by the inventors that various changes may be made in the peptide sequences of the disclosed compositions, or corresponding DNA sequences which encode said peptides without appreciable loss of their biological utility or activity.

5

TABLE 1

Amino Acids		Codons						
Alanine	Ala	A	GCA	GCC	GCG	GCU		
Cysteine	Cys	C	UGC	UGU				
Aspartic acid	Asp	D	GAC	GAU				
Glutamic acid	Glu	E	GAA	GAG				
Phenylalanine	Phe	F	UUC	UUU				
Glycine	Gly	G	GGA	GGC	GGG	GGU		
Histidine	His	H	CAC	CAU				
Isoleucine	Ile	I	AUA	AUC	AUU			
Lysine	Lys	K	AAA	AAG				
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU
Methionine	Met	M	AUG					
Asparagine	Asn	N	AAC	AAU				
Proline	Pro	P	CCA	CCC	CCG	CCU		
Glutamine	Gln	Q	CAA	CAG				
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU
Threonine	Thr	T	ACA	ACC	ACG	ACU		
Valine	Val	V	GUA	GUC	GUG	GUU		
Tryptophan	Trp	W	UGG					
Tyrosine	Tyr	Y	UAC	UAU				

In making such changes, the hydrophobic index of amino acids may be considered. The importance of the hydrophobic amino acid index in conferring interactive biologic function on a protein is generally understood in the art (Kyte and 10 Doolittle, 1982, incorporated herein by reference). It is accepted that the relative

Doolittle, 1982, incorporated herein by reference). It is accepted that the relative hydrophobic character of the amino acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of the protein with other molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and 5 the like. Each amino acid has been assigned a hydrophobic index on the basis of its hydrophobicity and charge characteristics (Kyte and Doolittle, 1982). These values are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); 10 glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

It is known in the art that certain amino acids may be substituted by other amino acids having a similar hydrophobic index or score and still result in a protein with similar biological activity, *i.e.* still obtain a biological functionally equivalent protein. In making such changes, the substitution of amino acids whose 15 hydrophobic indices are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred. It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U. S. Patent 4,554,101 (specifically incorporated herein by reference in its entirety), states that the greatest local average hydrophilicity of a 20 protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein.

As detailed in U. S. Patent 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate ($+3.0 \pm 1$); glutamate ($+3.0 \pm 1$); serine (+0.3); asparagine (+0.2); glutamine 25 (+0.2); glycine (0); threonine (-0.4); proline (-0.5 ± 1); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); tryptophan (-3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent protein. In 30 such changes, the substitution of amino acids whose hydrophilicity values are within ± 2

is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred.

As outlined above, amino acid substitutions are generally therefore based on the relative similarity of the amino acid side-chain substituents, for example, their 5 hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions that take various of the foregoing characteristics into consideration are well known to those of skill in the art and include: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine and asparagine; and valine, leucine and isoleucine.

In addition, any polynucleotide may be further modified to increase 10 stability *in vivo*. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends; the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages in the backbone; and/or the inclusion of nontraditional bases such as inosine, queosine and wybutosine, as well as acetyl-methyl-, thio- and other modified forms of adenine, cytidine, guanine, thymine and 15 uridine.

IN VIVO POLYNUCLEOTIDE DELIVERY TECHNIQUES

In additional embodiments, genetic constructs comprising one or more of the polynucleotides of the invention are introduced into cells *in vivo*. This may be achieved using any of a variety of well known approaches, several of which are outlined 20 below for the purpose of illustration.

1. ADENOVIRUS

One of the preferred methods for *in vivo* delivery of one or more nucleic acid sequences involves the use of an adenovirus expression vector. "Adenovirus expression vector" is meant to include those constructs containing adenovirus sequences 25 sufficient to (a) support packaging of the construct and (b) to express a polynucleotide that has been cloned therein in a sense or antisense orientation. Of course, in the context of an antisense construct, expression does not require that the gene product be synthesized.

The expression vector comprises a genetically engineered form of an adenovirus. Knowledge of the genetic organization of adenovirus, a 36 kb, linear, double-stranded DNA virus, allows substitution of large pieces of adenoviral DNA with foreign sequences up to 7 kb (Grunhaus and Horwitz, 1992). In contrast to retrovirus, 5 the adenoviral infection of host cells does not result in chromosomal integration because adenoviral DNA can replicate in an episomal manner without potential genotoxicity. Also, adenoviruses are structurally stable, and no genome rearrangement has been detected after extensive amplification. Adenovirus can infect virtually all epithelial cells regardless of their cell cycle stage. So far, adenoviral infection appears to be 10 linked only to mild disease such as acute respiratory disease in humans.

Adenovirus is particularly suitable for use as a gene transfer vector because of its mid-sized genome, ease of manipulation, high titer, wide target-cell range and high infectivity. Both ends of the viral genome contain 100-200 base pair inverted repeats (ITRs), which are *cis* elements necessary for viral DNA replication and 15 packaging. The early (E) and late (L) regions of the genome contain different transcription units that are divided by the onset of viral DNA replication. The E1 region (E1A and E1B) encodes proteins responsible for the regulation of transcription of the viral genome and a few cellular genes. The expression of the E2 region (E2A and E2B) results in the synthesis of the proteins for viral DNA replication. These proteins are 20 involved in DNA replication, late gene expression and host cell shut-off (Renan, 1990). The products of the late genes, including the majority of the viral capsid proteins, are expressed only after significant processing of a single primary transcript issued by the major late promoter (MLP). The MLP, (located at 16.8 m.u.) is particularly efficient during the late phase of infection, and all the mRNA's issued from this promoter possess 25 a 5'-tripartite leader (TPL) sequence which makes them preferred mRNA's for translation.

In a current system, recombinant adenovirus is generated from homologous recombination between shuttle vector and provirus vector. Due to the possible recombination between two proviral vectors, wild-type adenovirus may be 30 generated from this process. Therefore, it is critical to isolate a single clone of virus from an individual plaque and examine its genomic structure.

Generation and propagation of the current adenovirus vectors, which are replication deficient, depend on a unique helper cell line, designated 293, which was transformed from human embryonic kidney cells by Ad5 DNA fragments and constitutively expresses E1 proteins (Graham *et al.*, 1977). Since the E3 region is 5 dispensable from the adenovirus genome (Jones and Shenk, 1978), the current adenovirus vectors, with the help of 293 cells, carry foreign DNA in either the E1, the D3 or both regions (Graham and Prevec, 1991). In nature, adenovirus can package approximately 105% of the wild-type genome (Ghosh-Choudhury *et al.*, 1987), providing capacity for about 2 extra kB of DNA. Combined with the approximately 5.5 10 kB of DNA that is replaceable in the E1 and E3 regions, the maximum capacity of the current adenovirus vector is under 7.5 kB, or about 15% of the total length of the vector. More than 80% of the adenovirus viral genome remains in the vector backbone and is 15 the source of vector-borne cytotoxicity. Also, the replication deficiency of the E1-deleted virus is incomplete. For example, leakage of viral gene expression has been observed with the currently available vectors at high multiplicities of infection (MOI) (Mulligan, 1993).

Helper cell lines may be derived from human cells such as human embryonic kidney cells, muscle cells, hematopoietic cells or other human embryonic mesenchymal or epithelial cells. Alternatively, the helper cells may be derived from the 20 cells of other mammalian species that are permissive for human adenovirus. Such cells include, *e.g.*, Vero cells or other monkey embryonic mesenchymal or epithelial cells. As stated above, the currently preferred helper cell line is 293.

Recently, Racher *et al.* (1995) disclosed improved methods for culturing 293 cells and propagating adenovirus. In one format, natural cell aggregates are grown 25 by inoculating individual cells into 1 liter siliconized spinner flasks (Techne, Cambridge, UK) containing 100-200 ml of medium. Following stirring at 40 rpm, the cell viability is estimated with trypan blue. In another format, Fibra-Cel microcarriers (Bibby Sterlin, Stone, UK) (5 g/l) is employed as follows. A cell inoculum, resuspended in 5 ml of medium, is added to the carrier (50 ml) in a 250 ml Erlenmeyer 30 flask and left stationary, with occasional agitation, for 1 to 4 h. The medium is then replaced with 50 ml of fresh medium and shaking initiated. For virus production, cells

are allowed to grow to about 80% confluence, after which time the medium is replaced (to 25% of the final volume) and adenovirus added at an MOI of 0.05. Cultures are left stationary overnight, following which the volume is increased to 100% and shaking commenced for another 72 h.

5 Other than the requirement that the adenovirus vector be replication defective, or at least conditionally defective, the nature of the adenovirus vector is not believed to be crucial to the successful practice of the invention. The adenovirus may be of any of the 42 different known serotypes or subgroups A-F. Adenovirus type 5 of subgroup C is the preferred starting material in order to obtain a conditional replication-
10 defective adenovirus vector for use in the present invention, since Adenovirus type 5 is a human adenovirus about which a great deal of biochemical and genetic information is known, and it has historically been used for most constructions employing adenovirus as a vector.

15 As stated above, the typical vector according to the present invention is replication defective and will not have an adenovirus E1 region. Thus, it will be most convenient to introduce the polynucleotide encoding the gene of interest at the position from which the E1-coding sequences have been removed. However, the position of insertion of the construct within the adenovirus sequences is not critical to the invention. The polynucleotide encoding the gene of interest may also be inserted in lieu
20 of the deleted E3 region in E3 replacement vectors as described by Karlsson *et al.* (1986) or in the E4 region where a helper cell line or helper virus complements the E4 defect.

25 Adenovirus is easy to grow and manipulate and exhibits broad host range *in vitro* and *in vivo*. This group of viruses can be obtained in high titers, *e.g.*, 10^9 - 10^{11} plaque-forming units per ml, and they are highly infective. The life cycle of adenovirus does not require integration into the host cell genome. The foreign genes delivered by adenovirus vectors are episomal and, therefore, have low genotoxicity to host cells. No side effects have been reported in studies of vaccination with wild-type adenovirus (Couch *et al.*, 1963; Top *et al.*, 1971), demonstrating their safety and therapeutic
30 potential as *in vivo* gene transfer vectors.

Adenovirus vectors have been used in eukaryotic gene expression (Levrero *et al.*, 1991; Gomez-Foix *et al.*, 1992) and vaccine development (Grunhaus and Horwitz, 1992; Graham and Prevec, 1992). Recently, animal studies suggested that recombinant adenovirus could be used for gene therapy (Stratford-Perricaudet and 5 Perricaudet, 1991; Stratford-Perricaudet *et al.*, 1990; Rich *et al.*, 1993). Studies in administering recombinant adenovirus to different tissues include trachea instillation (Rosenfeld *et al.*, 1991; Rosenfeld *et al.*, 1992), muscle injection (Ragot *et al.*, 1993), peripheral intravenous injections (Herz and Gerard, 1993) and stereotactic inoculation into the brain (Le Gal La Salle *et al.*, 1993).

10 2. RETROVIRUSES

The retroviruses are a group of single-stranded RNA viruses characterized by an ability to convert their RNA to double-stranded DNA in infected cells by a process of reverse-transcription (Coffin, 1990). The resulting DNA then stably integrates into cellular chromosomes as a provirus and directs synthesis of viral 15 proteins. The integration results in the retention of the viral gene sequences in the recipient cell and its descendants. The retroviral genome contains three genes, gag, pol, and env that code for capsid proteins, polymerase enzyme, and envelope components, respectively. A sequence found upstream from the gag gene contains a signal for packaging of the genome into virions. Two long terminal repeat (LTR) sequences are 20 present at the 5' and 3' ends of the viral genome. These contain strong promoter and enhancer sequences and are also required for integration in the host cell genome (Coffin, 1990).

In order to construct a retroviral vector, a nucleic acid encoding one or 25 more oligonucleotide or polynucleotide sequences of interest is inserted into the viral genome in the place of certain viral sequences to produce a virus that is replication-defective. In order to produce virions, a packaging cell line containing the gag, pol, and env genes but without the LTR and packaging components is constructed (Mann *et al.*, 1983). When a recombinant plasmid containing a cDNA, together with the retroviral LTR and packaging sequences is introduced into this cell line (by calcium phosphate 30 precipitation for example), the packaging sequence allows the RNA transcript of the

recombinant plasmid to be packaged into viral particles, which are then secreted into the culture media (Nicolas and Rubenstein, 1988; Temin, 1986; Mann *et al.*, 1983). The media containing the recombinant retroviruses is then collected, optionally concentrated, and used for gene transfer. Retroviral vectors are able to infect a broad 5 variety of cell types. However, integration and stable expression require the division of host cells (Paskind *et al.*, 1975).

A novel approach designed to allow specific targeting of retrovirus vectors was recently developed based on the chemical modification of a retrovirus by the chemical addition of lactose residues to the viral envelope. This modification could 10 permit the specific infection of hepatocytes *via* sialoglycoprotein receptors.

A different approach to targeting of recombinant retroviruses was designed in which biotinylated antibodies against a retroviral envelope protein and against a specific cell receptor were used. The antibodies were coupled *via* the biotin components by using streptavidin (Roux *et al.*, 1989). Using antibodies against major 15 histocompatibility complex class I and class II antigens, they demonstrated the infection of a variety of human cells that bore those surface antigens with an ecotropic virus *in vitro* (Roux *et al.*, 1989).

3. ADENO-ASSOCIATED VIRUSES

AAV (Ridgeway, 1988; Hermonat and Muzyczka, 1984) is a parovirus, 20 discovered as a contamination of adenoviral stocks. It is a ubiquitous virus (antibodies are present in 85% of the US human population) that has not been linked to any disease. It is also classified as a dependovirus, because its replication is dependent on the presence of a helper virus, such as adenovirus. Five serotypes have been isolated, of 25 which AAV-2 is the best characterized. AAV has a single-stranded linear DNA that is encapsidated into capsid proteins VP1, VP2 and VP3 to form an icosahedral virion of 20 to 24 nm in diameter (Muzyczka and McLaughlin, 1988).

The AAV DNA is approximately 4.7 kilobases long. It contains two open reading frames and is flanked by two ITRs. There are two major genes in the AAV genome: *rep* and *cap*. The *rep* gene codes for proteins responsible for viral 30 replications, whereas *cap* codes for capsid protein VP1-3. Each ITR forms a T-shaped

hairpin structure. These terminal repeats are the only essential *cis* components of the AAV for chromosomal integration. Therefore, the AAV can be used as a vector with all viral coding sequences removed and replaced by the cassette of genes for delivery. Three viral promoters have been identified and named p5, p19, and p40, according to 5 their map position. Transcription from p5 and p19 results in production of rep proteins, and transcription from p40 produces the capsid proteins (Hermonat and Muzyczka, 1984).

There are several factors that prompted researchers to study the possibility of using rAAV as an expression vector. One is that the requirements for 10 delivering a gene to integrate into the host chromosome are surprisingly few. It is necessary to have the 145-bp ITRs, which are only 6% of the AAV genome. This leaves room in the vector to assemble a 4.5-kb DNA insertion. While this carrying capacity may prevent the AAV from delivering large genes, it is amply suited for delivering the antisense constructs of the present invention.

15 AAV is also a good choice of delivery vehicles due to its safety. There is a relatively complicated rescue mechanism: not only wild type adenovirus but also AAV genes are required to mobilize rAAV. Likewise, AAV is not pathogenic and not associated with any disease. The removal of viral coding sequences minimizes immune reactions to viral gene expression, and therefore, rAAV does not evoke an inflammatory 20 response.

4. OTHER VIRAL VECTORS AS EXPRESSION CONSTRUCTS

Other viral vectors may be employed as expression constructs in the present invention for the delivery of oligonucleotide or polynucleotide sequences to a host cell. Vectors derived from viruses such as vaccinia virus (Ridgeway, 1988; Coupar 25 *et al.*, 1988), lentiviruses, polio viruses and herpes viruses may be employed. They offer several attractive features for various mammalian cells (Friedmann, 1989; Ridgeway, 1988; Coupar *et al.*, 1988; Horwitz *et al.*, 1990).

With the recent recognition of defective hepatitis B viruses, new insight was gained into the structure-function relationship of different viral sequences. *In vitro* 30 studies showed that the virus could retain the ability for helper-dependent packaging

and reverse transcription despite the deletion of up to 80% of its genome (Horwitz *et al.*, 1990). This suggested that large portions of the genome could be replaced with foreign genetic material. The hepatotropism and persistence (integration) were particularly attractive properties for liver-directed gene transfer. Chang *et al.* (1991) 5 introduced the chloramphenicol acetyltransferase (CAT) gene into duck hepatitis B virus genome in the place of the polymerase, surface, and pre-surface coding sequences. It was cotransfected with wild-type virus into an avian hepatoma cell line. Culture media containing high titers of the recombinant virus were used to infect primary duckling hepatocytes. Stable CAT gene expression was detected for at least 24 days 10 after transfection (Chang *et al.*, 1991).

5. NON-VIRAL VECTORS

In order to effect expression of the oligonucleotide or polynucleotide sequences of the present invention, the expression construct must be delivered into a cell. This delivery may be accomplished *in vitro*, as in laboratory procedures for 15 transforming cells lines, or *in vivo* or *ex vivo*, as in the treatment of certain disease states. As described above, one preferred mechanism for delivery is *via* viral infection where the expression construct is encapsulated in an infectious viral particle.

Once the expression construct has been delivered into the cell the nucleic acid encoding the desired oligonucleotide or polynucleotide sequences may be 20 positioned and expressed at different sites. In certain embodiments, the nucleic acid encoding the construct may be stably integrated into the genome of the cell. This integration may be in the specific location and orientation *via* homologous recombination (gene replacement) or it may be integrated in a random, non-specific location (gene augmentation). In yet further embodiments, the nucleic acid may be 25 stably maintained in the cell as a separate, episomal segment of DNA. Such nucleic acid segments or "episomes" encode sequences sufficient to permit maintenance and replication independent of or in synchronization with the host cell cycle. How the expression construct is delivered to a cell and where in the cell the nucleic acid remains is dependent on the type of expression construct employed.

In certain embodiments of the invention, the expression construct comprising one or more oligonucleotide or polynucleotide sequences may simply consist of naked recombinant DNA or plasmids. Transfer of the construct may be performed by any of the methods mentioned above which physically or chemically 5 permeabilize the cell membrane. This is particularly applicable for transfer *in vitro* but it may be applied to *in vivo* use as well. Dubensky *et al.* (1984) successfully injected polyomavirus DNA in the form of calcium phosphate precipitates into liver and spleen of adult and newborn mice demonstrating active viral replication and acute infection. Benvenisty and Reshef (1986) also demonstrated that direct intraperitoneal injection of 10 calcium phosphate-precipitated plasmids results in expression of the transfected genes. It is envisioned that DNA encoding a gene of interest may also be transferred in a similar manner *in vivo* and express the gene product.

Another embodiment of the invention for transferring a naked DNA expression construct into cells may involve particle bombardment. This method 15 depends on the ability to accelerate DNA-coated microprojectiles to a high velocity allowing them to pierce cell membranes and enter cells without killing them (Klein *et al.*, 1987). Several devices for accelerating small particles have been developed. One such device relies on a high voltage discharge to generate an electrical current, which in turn provides the motive force (Yang *et al.*, 1990). The microprojectiles used have 20 consisted of biologically inert substances such as tungsten or gold beads.

Selected organs including the liver, skin, and muscle tissue of rats and mice have been bombarded *in vivo* (Yang *et al.*, 1990; Zelenin *et al.*, 1991). This may require surgical exposure of the tissue or cells, to eliminate any intervening tissue between the gun and the target organ, *i.e. ex vivo* treatment. Again, DNA encoding a 25 particular gene may be delivered *via* this method and still be incorporated by the present invention.

ANTISENSE OLIGONUCLEOTIDES

The end result of the flow of genetic information is the synthesis of protein. DNA is transcribed by polymerases into messenger RNA and translated on the 30 ribosome to yield a folded, functional protein. Thus there are several steps along the

route where protein synthesis can be inhibited. The native DNA segment coding for a polypeptide described herein, as all such mammalian DNA strands, has two strands: a sense strand and an antisense strand held together by hydrogen bonding. The messenger RNA coding for polypeptide has the same nucleotide sequence as the sense 5 DNA strand except that the DNA thymidine is replaced by uridine. Thus, synthetic antisense nucleotide sequences will bind to a mRNA and inhibit expression of the protein encoded by that mRNA.

The targeting of antisense oligonucleotides to mRNA is thus one mechanism to shut down protein synthesis, and, consequently, represents a powerful 10 and targeted therapeutic approach. For example, the synthesis of polygalacturonase and the muscarine type 2 acetylcholine receptor are inhibited by antisense oligonucleotides directed to their respective mRNA sequences (U. S. Patent 5,739,119 and U. S. Patent 5,759,829, each specifically incorporated herein by reference in its entirety). Further, examples of antisense inhibition have been demonstrated with the 15 nuclear protein cyclin, the multiple drug resistance gene (MDG1), ICAM-1, E-selectin, STK-1, striatal GABA_A receptor and human EGF (Jaskulski *et al.*, 1988; Vasanthakumar and Ahmed, 1989; Peris *et al.*, 1998; U. S. Patent 5,801,154; U. S. Patent 5,789,573; U. S. Patent 5,718,709 and U. S. Patent 5,610,288, each specifically incorporated herein by reference in its entirety). Antisense constructs have also been 20 described that inhibit and can be used to treat a variety of abnormal cellular proliferations, *e.g.* cancer (U. S. Patent 5,747,470; U. S. Patent 5,591,317 and U. S. Patent 5,783,683, each specifically incorporated herein by reference in its entirety).

Therefore, in exemplary embodiments, the invention provides 25 oligonucleotide sequences that comprise all, or a portion of, any sequence that is capable of specifically binding to polynucleotide sequence described herein, or a complement thereof. In one embodiment, the antisense oligonucleotides comprise DNA or derivatives thereof. In another embodiment, the oligonucleotides comprise RNA or derivatives thereof. In a third embodiment, the oligonucleotides are modified DNAs comprising a phosphorothioated modified backbone. In a fourth embodiment, the 30 oligonucleotide sequences comprise peptide nucleic acids or derivatives thereof. In each case, preferred compositions comprise a sequence region that is complementary,

and more preferably substantially-complementary, and even more preferably, completely complementary to one or more portions of polynucleotides disclosed herein.

Selection of antisense compositions specific for a given gene sequence is based upon analysis of the chosen target sequence (*i.e.* in these illustrative examples the 5 rat and human sequences) and determination of secondary structure, T_m , binding energy, relative stability, and antisense compositions were selected based upon their relative inability to form dimers, hairpins, or other secondary structures that would reduce or prohibit specific binding to the target mRNA in a host cell.

Highly preferred target regions of the mRNA, are those which are at or 10 near the AUG translation initiation codon, and those sequences which were substantially complementary to 5' regions of the mRNA. These secondary structure analyses and target site selection considerations were performed using v.4 of the OLIGO primer analysis software (Rychlik, 1997) and the BLASTN 2.0.5 algorithm software (Altschul *et al.*, 1997).

15 The use of an antisense delivery method employing a short peptide vector, termed MPG (27 residues), is also contemplated. The MPG peptide contains a hydrophobic domain derived from the fusion sequence of HIV gp41 and a hydrophilic domain from the nuclear localization sequence of SV40 T-antigen (Morris *et al.*, 1997). It has been demonstrated that several molecules of the MPG peptide coat the antisense 20 oligonucleotides and can be delivered into cultured mammalian cells in less than 1 hour with relatively high efficiency (90%). Further, the interaction with MPG strongly increases both the stability of the oligonucleotide to nuclease and the ability to cross the plasma membrane (Morris *et al.*, 1997).

RIBOZYMES

25 Although proteins traditionally have been used for catalysis of nucleic acids, another class of macromolecules has emerged as useful in this endeavor. Ribozymes are RNA-protein complexes that cleave nucleic acids in a site-specific fashion. Ribozymes have specific catalytic domains that possess endonuclease activity (Kim and Cech, 1987; Gerlach *et al.*, 1987; Forster and Symons, 1987). For example, a 30 large number of ribozymes accelerate phosphoester transfer reactions with a high degree

of specificity, often cleaving only one of several phosphoesters in an oligonucleotide substrate (Cech *et al.*, 1981; Michel and Westhof, 1990; Reinhold-Hurek and Shub, 1992). This specificity has been attributed to the requirement that the substrate bind via specific base-pairing interactions to the internal guide sequence ("IGS") of the ribozyme 5 prior to chemical reaction.

Ribozyme catalysis has primarily been observed as part of sequence-specific cleavage/ligation reactions involving nucleic acids (Joyce, 1989; Cech *et al.*, 1981). For example, U. S. Patent No. 5,354,855 (specifically incorporated herein by reference) reports that certain ribozymes can act as endonucleases with a sequence 10 specificity greater than that of known ribonucleases and approaching that of the DNA restriction enzymes. Thus, sequence-specific ribozyme-mediated inhibition of gene expression may be particularly suited to therapeutic applications (Scanlon *et al.*, 1991; Sarver *et al.*, 1990). Recently, it was reported that ribozymes elicited genetic changes 15 in some cells lines to which they were applied; the altered genes included the oncogenes H-ras, c-fos and genes of HIV. Most of this work involved the modification of a target mRNA, based on a specific mutant codon that is cleaved by a specific ribozyme.

Six basic varieties of naturally-occurring enzymatic RNAs are known presently. Each can catalyze the hydrolysis of RNA phosphodiester bonds *in trans* (and thus can cleave other RNA molecules) under physiological conditions. In general, 20 enzymatic nucleic acids act by first binding to a target RNA. Such binding occurs through the target binding portion of a enzymatic nucleic acid which is held in close proximity to an enzymatic portion of the molecule that acts to cleave the target RNA. Thus, the enzymatic nucleic acid first recognizes and then binds a target RNA through 25 complementary base-pairing, and once bound to the correct site, acts enzymatically to cut the target RNA. Strategic cleavage of such a target RNA will destroy its ability to direct synthesis of an encoded protein. After an enzymatic nucleic acid has bound and cleaved its RNA target, it is released from that RNA to search for another target and can repeatedly bind and cleave new targets.

The enzymatic nature of a ribozyme is advantageous over many 30 technologies, such as antisense technology (where a nucleic acid molecule simply binds to a nucleic acid target to block its translation) since the concentration of ribozyme

necessary to affect a therapeutic treatment is lower than that of an antisense oligonucleotide. This advantage reflects the ability of the ribozyme to act enzymatically. Thus, a single ribozyme molecule is able to cleave many molecules of target RNA. In addition, the ribozyme is a highly specific inhibitor, with the specificity 5 of inhibition depending not only on the base pairing mechanism of binding to the target RNA, but also on the mechanism of target RNA cleavage. Single mismatches, or base- substitutions, near the site of cleavage can completely eliminate catalytic activity of a ribozyme. Similar mismatches in antisense molecules do not prevent their action (Woolf *et al.*, 1992). Thus, the specificity of action of a ribozyme is greater than that of 10 an antisense oligonucleotide binding the same RNA site.

The enzymatic nucleic acid molecule may be formed in a hammerhead, hairpin, a hepatitis δ virus, group I intron or RNaseP RNA (in association with an RNA guide sequence) or Neurospora VS RNA motif. Examples of hammerhead motifs are described by Rossi *et al.* (1992). Examples of hairpin motifs are described by Hampel 15 *et al.* (Eur. Pat. Appl. Publ. No. EP 0360257), Hampel and Tritz (1989), Hampel *et al.* (1990) and U. S. Patent 5,631,359 (specifically incorporated herein by reference). An example of the hepatitis δ virus motif is described by Perrotta and Been (1992); an example of the RNaseP motif is described by Guerrier-Takada *et al.* (1983); Neurospora VS RNA ribozyme motif is described by Collins (Saville and Collins, 1990; Saville and 20 Collins, 1991; Collins and Olive, 1993); and an example of the Group I intron is described in (U. S. Patent 4,987,071, specifically incorporated herein by reference). All that is important in an enzymatic nucleic acid molecule of this invention is that it has a specific substrate binding site which is complementary to one or more of the target gene RNA regions, and that it have nucleotide sequences within or surrounding that substrate 25 binding site which impart an RNA cleaving activity to the molecule. Thus the ribozyme constructs need not be limited to specific motifs mentioned herein.

In certain embodiments, it may be important to produce enzymatic cleaving agents which exhibit a high degree of specificity for the RNA of a desired target, such as one of the sequences disclosed herein. The enzymatic nucleic acid 30 molecule is preferably targeted to a highly conserved sequence region of a target mRNA. Such enzymatic nucleic acid molecules can be delivered exogenously to

specific cells as required. Alternatively, the ribozymes can be expressed from DNA or RNA vectors that are delivered to specific cells.

Small enzymatic nucleic acid motifs (e.g., of the hammerhead or the hairpin structure) may also be used for exogenous delivery. The simple structure of 5 these molecules increases the ability of the enzymatic nucleic acid to invade targeted regions of the mRNA structure. Alternatively, catalytic RNA molecules can be expressed within cells from eukaryotic promoters (e.g., Scanlon *et al.*, 1991; Kashani-Sabet *et al.*, 1992; Dropulic *et al.*, 1992; Weerasinghe *et al.*, 1991; Ojwang *et al.*, 1992; Chen *et al.*, 1992; Sarver *et al.*, 1990). Those skilled in the art realize that any ribozyme 10 can be expressed in eukaryotic cells from the appropriate DNA vector. The activity of such ribozymes can be augmented by their release from the primary transcript by a second ribozyme (Int. Pat. Appl. Publ. No. WO 93/23569, and Int. Pat. Appl. Publ. No. WO 94/02595, both hereby incorporated by reference; Ohkawa *et al.*, 1992; Taira *et al.*, 1991; and Ventura *et al.*, 1993).

15 Ribozymes may be added directly, or can be complexed with cationic lipids, lipid complexes, packaged within liposomes, or otherwise delivered to target cells. The RNA or RNA complexes can be locally administered to relevant tissues *ex vivo*, or *in vivo* through injection, aerosol inhalation, infusion pump or stent, with or without their incorporation in biopolymers.

20 Ribozymes may be designed as described in Int. Pat. Appl. Publ. No. WO 93/23569 and Int. Pat. Appl. Publ. No. WO 94/02595, each specifically incorporated herein by reference) and synthesized to be tested *in vitro* and *in vivo*, as described. Such ribozymes can also be optimized for delivery. While specific examples are provided, those in the art will recognize that equivalent RNA targets in 25 other species can be utilized when necessary.

Hammerhead or hairpin ribozymes may be individually analyzed by computer folding (Jaeger *et al.*, 1989) to assess whether the ribozyme sequences fold into the appropriate secondary structure. Those ribozymes with unfavorable intramolecular interactions between the binding arms and the catalytic core are 30 eliminated from consideration. Varying binding arm lengths can be chosen to optimize

activity. Generally, at least 5 or so bases on each arm are able to bind to, or otherwise interact with, the target RNA.

Ribozymes of the hammerhead or hairpin motif may be designed to anneal to various sites in the mRNA message, and can be chemically synthesized. The 5 method of synthesis used follows the procedure for normal RNA synthesis as described in Usman *et al.* (1987) and in Scaringe *et al.* (1990) and makes use of common nucleic acid protecting and coupling groups, such as dimethoxytrityl at the 5'-end, and phosphoramidites at the 3'-end. Average stepwise coupling yields are typically >98%. Hairpin ribozymes may be synthesized in two parts and annealed to reconstruct an 10 active ribozyme (Chowrira and Burke, 1992). Ribozymes may be modified extensively to enhance stability by modification with nuclease resistant groups, for example, 2'-amino, 2'-C-allyl, 2'-flouro, 2'-o-methyl, 2'-H (for a review see e.g., Usman and Cedergren, 1992). Ribozymes may be purified by gel electrophoresis using general methods or by high pressure liquid chromatography and resuspended in water.

15 Ribozyme activity can be optimized by altering the length of the ribozyme binding arms, or chemically synthesizing ribozymes with modifications that prevent their degradation by serum ribonucleases (see e.g., Int. Pat. Appl. Publ. No. WO 92/07065; Perrault *et al.*, 1990; Pieken *et al.*, 1991; Usman and Cedergren, 1992; Int. Pat. Appl. Publ. No. WO 93/15187; Int. Pat. Appl. Publ. No. WO 91/03162; Eur. Pat. 20 Appl. Publ. No. 92110298.4; U. S. Patent 5,334,711; and Int. Pat. Appl. Publ. No. WO 94/13688, which describe various chemical modifications that can be made to the sugar 30 moieties of enzymatic RNA molecules), modifications which enhance their efficacy in cells, and removal of stem II bases to shorten RNA synthesis times and reduce chemical requirements.

25 Sullivan *et al.* (Int. Pat. Appl. Publ. No. WO 94/02595) describes the general methods for delivery of enzymatic RNA molecules. Ribozymes may be administered to cells by a variety of methods known to those familiar to the art, including, but not restricted to, encapsulation in liposomes, by iontophoresis, or by incorporation into other vehicles, such as hydrogels, cyclodextrins, biodegradable 30 nanocapsules, and bioadhesive microspheres. For some indications, ribozymes may be directly delivered *ex vivo* to cells or tissues with or without the aforementioned vehicles.

Alternatively, the RNA/vehicle combination may be locally delivered by direct inhalation, by direct injection or by use of a catheter, infusion pump or stent. Other routes of delivery include, but are not limited to, intravascular, intramuscular, subcutaneous or joint injection, aerosol inhalation, oral (tablet or pill form), topical, 5 systemic, ocular, intraperitoneal and/or intrathecal delivery. More detailed descriptions of ribozyme delivery and administration are provided in Int. Pat. Appl. Publ. No. WO 94/02595 and Int. Pat. Appl. Publ. No. WO 93/23569, each specifically incorporated herein by reference.

Another means of accumulating high concentrations of a ribozyme(s) 10 within cells is to incorporate the ribozyme-encoding sequences into a DNA expression vector. Transcription of the ribozyme sequences are driven from a promoter for eukaryotic RNA polymerase I (pol I), RNA polymerase II (pol II), or RNA polymerase III (pol III). Transcripts from pol II or pol III promoters will be expressed at high levels in all cells; the levels of a given pol II promoter in a given cell type will depend on the 15 nature of the gene regulatory sequences (enhancers, silencers, *etc.*) present nearby. Prokaryotic RNA polymerase promoters may also be used, providing that the prokaryotic RNA polymerase enzyme is expressed in the appropriate cells (Elroy-Stein and Moss, 1990; Gao and Huang, 1993; Lieber *et al.*, 1993; Zhou *et al.*, 1990). Ribozymes expressed from such promoters can function in mammalian cells (*e.g.* 20 Kashani-Saber *et al.*, 1992; Ojwang *et al.*, 1992; Chen *et al.*, 1992; Yu *et al.*, 1993; L'Huillier *et al.*, 1992; Lisziewicz *et al.*, 1993). Such transcription units can be incorporated into a variety of vectors for introduction into mammalian cells, including 25 but not restricted to, plasmid DNA vectors, viral DNA vectors (such as adenovirus or adeno-associated vectors), or viral RNA vectors (such as retroviral, semliki forest virus, sindbis virus vectors).

Ribozymes may be used as diagnostic tools to examine genetic drift and mutations within diseased cells. They can also be used to assess levels of the target RNA molecule. The close relationship between ribozyme activity and the structure of the target RNA allows the detection of mutations in any region of the molecule which 30 alters the base-pairing and three-dimensional structure of the target RNA. By using multiple ribozymes, one may map nucleotide changes which are important to RNA

structure and function *in vitro*, as well as in cells and tissues. Cleavage of target RNAs with ribozymes may be used to inhibit gene expression and define the role (essentially) of specified gene products in the progression of disease. In this manner, other genetic targets may be defined as important mediators of the disease. These studies will lead to

5 better treatment of the disease progression by affording the possibility of combinational therapies (e.g., multiple ribozymes targeted to different genes, ribozymes coupled with known small molecule inhibitors, or intermittent treatment with combinations of ribozymes and/or other chemical or biological molecules). Other *in vitro* uses of ribozymes are well known in the art, and include detection of the presence of mRNA

10 associated with an IL-5 related condition. Such RNA is detected by determining the presence of a cleavage product after treatment with a ribozyme using standard methodology.

PEPTIDE NUCLEIC ACIDS

In certain embodiments, the inventors contemplate the use of peptide

15 nucleic acids (PNAs) in the practice of the methods of the invention. PNA is a DNA mimic in which the nucleobases are attached to a pseudopeptide backbone (Good and Nielsen, 1997). PNA is able to be utilized in a number methods that traditionally have used RNA or DNA. Often PNA sequences perform better in techniques than the corresponding RNA or DNA sequences and have utilities that are not inherent to RNA

20 or DNA. A review of PNA including methods of making, characteristics of, and methods of using, is provided by Corey (1997) and is incorporated herein by reference. As such, in certain embodiments, one may prepare PNA sequences that are complementary to one or more portions of the ACE mRNA sequence, and such PNA compositions may be used to regulate, alter, decrease, or reduce the translation of ACE-

25 specific mRNA, and thereby alter the level of ACE activity in a host cell to which such PNA compositions have been administered.

PNAs have 2-aminoethyl-glycine linkages replacing the normal phosphodiester backbone of DNA (Nielsen *et al.*, 1991; Hanvey *et al.*, 1992; Hyrup and Nielsen, 1996; Nielsen, 1996). This chemistry has three important consequences:

30 firstly, in contrast to DNA or phosphorothioate oligonucleotides, PNAs are neutral

molecules; secondly, PNAs are achiral, which avoids the need to develop a stereoselective synthesis; and thirdly, PNA synthesis uses standard Boc (Dueholm *et al.*, 1994) or Fmoc (Thomson *et al.*, 1995) protocols for solid-phase peptide synthesis, although other methods, including a modified Merrifield method, have been used
5 (Christensen *et al.*, 1995).

PNA monomers or ready-made oligomers are commercially available from PerSeptive Biosystems (Framingham, MA). PNA syntheses by either Boc or Fmoc protocols are straightforward using manual or automated protocols (Norton *et al.*, 1995). The manual protocol lends itself to the production of chemically modified PNAs
10 or the simultaneous synthesis of families of closely related PNAs.

As with peptide synthesis, the success of a particular PNA synthesis will depend on the properties of the chosen sequence. For example, while in theory PNAs can incorporate any combination of nucleotide bases, the presence of adjacent purines can lead to deletions of one or more residues in the product. In expectation of this
15 difficulty, it is suggested that, in producing PNAs with adjacent purines, one should repeat the coupling of residues likely to be added inefficiently. This should be followed by the purification of PNAs by reverse-phase high-pressure liquid chromatography (Norton *et al.*, 1995) providing yields and purity of product similar to those observed during the synthesis of peptides.

20 Modifications of PNAs for a given application may be accomplished by coupling amino acids during solid-phase synthesis or by attaching compounds that contain a carboxylic acid group to the exposed N-terminal amine. Alternatively, PNAs can be modified after synthesis by coupling to an introduced lysine or cysteine. The ease with which PNAs can be modified facilitates optimization for better solubility or
25 for specific functional requirements. Once synthesized, the identity of PNAs and their derivatives can be confirmed by mass spectrometry. Several studies have made and utilized modifications of PNAs (Norton *et al.*, 1995; Haaima *et al.*, 1996; Stetsenko *et al.*, 1996; Petersen *et al.*, 1995; Ullmann *et al.*, 1996; Koch *et al.*, 1995; Orum *et al.*, 1995; Footer *et al.*, 1996; Griffith *et al.*, 1995; Kremsky *et al.*, 1996; Pardridge *et al.*,
30 1995; Boffa *et al.*, 1995; Landsdorp *et al.*, 1996; Gambacorti-Passerini *et al.*, 1996; Armitage *et al.*, 1997; Seeger *et al.*, 1997; Ruskowski *et al.*, 1997). U.S. Patent No.

5,700,922 discusses PNA-DNA-PNA chimeric molecules and their uses in diagnostics, modulating protein in organisms, and treatment of conditions susceptible to therapeutics.

In contrast to DNA and RNA, which contain negatively charged 5 linkages, the PNA backbone is neutral. In spite of this dramatic alteration, PNAs recognize complementary DNA and RNA by Watson-Crick pairing (Egholm *et al.*, 1993), validating the initial modeling by Nielsen *et al.* (1991). PNAs lack 3' to 5' polarity and can bind in either parallel or antiparallel fashion, with the antiparallel mode being preferred (Egholm *et al.*, 1993).

10 Hybridization of DNA oligonucleotides to DNA and RNA is destabilized by electrostatic repulsion between the negatively charged phosphate backbones of the complementary strands. By contrast, the absence of charge repulsion in PNA-DNA or PNA-RNA duplexes increases the melting temperature (T_m) and reduces the dependence of T_m on the concentration of mono- or divalent cations (Nielsen *et al.*, 1991). The 15 enhanced rate and affinity of hybridization are significant because they are responsible for the surprising ability of PNAs to perform strand invasion of complementary sequences within relaxed double-stranded DNA. In addition, the efficient hybridization at inverted repeats suggests that PNAs can recognize secondary structure effectively within double-stranded DNA. Enhanced recognition also occurs with PNAs 20 immobilized on surfaces, and Wang *et al.* have shown that support-bound PNAs can be used to detect hybridization events (Wang *et al.*, 1996).

One might expect that tight binding of PNAs to complementary sequences would also increase binding to similar (but not identical) sequences, reducing the sequence specificity of PNA recognition. As with DNA hybridization, however, 25 selective recognition can be achieved by balancing oligomer length and incubation temperature. Moreover, selective hybridization of PNAs is encouraged by PNA-DNA hybridization being less tolerant of base mismatches than DNA-DNA hybridization. For example, a single mismatch within a 16 bp PNA-DNA duplex can reduce the T_m by up to 15°C (Egholm *et al.*, 1993). This high level of discrimination has allowed the 30 development of several PNA-based strategies for the analysis of point mutations (Wang

et al., 1996; Carlsson *et al.*, 1996; Thiede *et al.*, 1996; Webb and Hurskainen, 1996; Perry-O'Keefe *et al.*, 1996).

High-affinity binding provides clear advantages for molecular recognition and the development of new applications for PNAs. For example, 11-13 5 nucleotide PNAs inhibit the activity of telomerase, a ribonucleo-protein that extends telomere ends using an essential RNA template, while the analogous DNA oligomers do not (Norton *et al.*, 1996).

Neutral PNAs are more hydrophobic than analogous DNA oligomers, and this can lead to difficulty solubilizing them at neutral pH, especially if the PNAs 10 have a high purine content or if they have the potential to form secondary structures. Their solubility can be enhanced by attaching one or more positive charges to the PNA termini (Nielsen *et al.*, 1991).

Findings by Allfrey and colleagues suggest that strand invasion will occur spontaneously at sequences within chromosomal DNA (Boffa *et al.*, 1995; Boffa 15 *et al.*, 1996). These studies targeted PNAs to triplet repeats of the nucleotides CAG and used this recognition to purify transcriptionally active DNA (Boffa *et al.*, 1995) and to inhibit transcription (Boffa *et al.*, 1996). This result suggests that if PNAs can be delivered within cells then they will have the potential to be general sequence-specific regulators of gene expression. Studies and reviews concerning the use of PNAs as 20 antisense and anti-gene agents include Nielsen *et al.* (1993b), Hanvey *et al.* (1992), and Good and Nielsen (1997). Koppelhus *et al.* (1997) have used PNAs to inhibit HIV-1 inverse transcription, showing that PNAs may be used for antiviral therapies.

Methods of characterizing the antisense binding properties of PNAs are discussed in Rose (1993) and Jensen *et al.* (1997). Rose uses capillary gel 25 electrophoresis to determine binding of PNAs to their complementary oligonucleotide, measuring the relative binding kinetics and stoichiometry. Similar types of measurements were made by Jensen *et al.* using BIACore™ technology.

Other applications of PNAs include use in DNA strand invasion (Nielsen *et al.*, 1991), antisense inhibition (Hanvey *et al.*, 1992), mutational analysis (Orum *et* 30 *al.*, 1993), enhancers of transcription (Mollegaard *et al.*, 1994), nucleic acid purification (Orum *et al.*, 1995), isolation of transcriptionally active genes (Boffa *et al.*, 1995),

blocking of transcription factor binding (Vickers *et al.*, 1995), genome cleavage (Veselkov *et al.*, 1996), biosensors (Wang *et al.*, 1996), *in situ* hybridization (Thisted *et al.*, 1996), and in an alternative to Southern blotting (Perry-O'Keefe, 1996).

POLYPEPTIDE COMPOSITIONS

5 The present invention, in other aspects, provides polypeptide compositions. Generally, a polypeptide of the invention will be an isolated polypeptide (or an epitope, variant, or active fragment thereof) derived from a mammalian species. Preferably, the polypeptide is encoded by a polynucleotide sequence disclosed herein or 10 a sequence which hybridizes under moderately stringent conditions to a polynucleotide sequence disclosed herein. Alternatively, the polypeptide may be defined as a polypeptide which comprises a contiguous amino acid sequence from an amino acid sequence disclosed herein, or which polypeptide comprises an entire amino acid sequence disclosed herein.

15 In the present invention, a polypeptide composition is also understood to comprise one or more polypeptides that are immunologically reactive with antibodies generated against a polypeptide of the invention, or to active fragments, or to variants or biological functional equivalents thereof.

20 Likewise, a polypeptide composition of the present invention is understood to comprise one or more polypeptides that are capable of eliciting antibodies that are immunologically reactive with one or more polypeptides encoded by one or more contiguous nucleic acid sequences contained in SEQ ID NO: 1-38, 42-204, 205, 207 and 210-290, or to active fragments, or to variants thereof, or to one or more nucleic acid sequences which hybridize to one or more of these sequences under conditions of moderate to high stringency.

25 As used herein, an active fragment of a polypeptide includes a whole or a portion of a polypeptide which is modified by conventional techniques, *e.g.*, mutagenesis, or by addition, deletion, or substitution, but which active fragment exhibits substantially the same structure function, antigenicity, etc., as a polypeptide as described herein.

In certain illustrative embodiments, the polypeptides of the invention will comprise at least an immunogenic portion of a breast tumor protein or a variant thereof, as described herein. As noted above, a "breast tumor protein" is a protein that is expressed by breast tumor cells. Proteins that are breast tumor proteins react 5 detectably within an immunoassay (such as an ELISA) with antisera from a patient with breast cancer. Polypeptides as described herein may be of any length. Additional sequences derived from the native protein and/or heterologous sequences may be present, and such sequences may (but need not) possess further immunogenic or antigenic properties.

10 An "immunogenic portion," as used herein is a portion of a protein that is recognized (*i.e.*, specifically bound) by a B-cell and/or T-cell surface antigen receptor. Such immunogenic portions generally comprise at least 5 amino acid residues, more preferably at least 10, and still more preferably at least 20 amino acid residues of a breast tumor protein or a variant thereof. Certain preferred immunogenic 15 portions include peptides in which an N-terminal leader sequence and/or transmembrane domain have been deleted. Other preferred immunogenic portions may contain a small N- and/or C-terminal deletion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids), relative to the mature protein.

10 Immunogenic portions may generally be identified using well known 20 techniques, such as those summarized in Paul, *Fundamental Immunology*, 3rd ed., 243-247 (Raven Press, 1993) and references cited therein. Such techniques include screening polypeptides for the ability to react with antigen-specific antibodies, antisera and/or T-cell lines or clones. As used herein, antisera and antibodies are "antigen-specific" if they specifically bind to an antigen (*i.e.*, they react with the protein in an 25 ELISA or other immunoassay, and do not react detectably with unrelated proteins). Such antisera and antibodies may be prepared as described herein, and using well known techniques. An immunogenic portion of a native breast tumor protein is a portion that reacts with such antisera and/or T-cells at a level that is not substantially less than the reactivity of the full length polypeptide (*e.g.*, in an ELISA and/or T-cell 30 reactivity assay). Such immunogenic portions may react within such assays at a level that is similar to or greater than the reactivity of the full length polypeptide. Such

screens may generally be performed using methods well known to those of ordinary skill in the art, such as those described in Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. For example, a polypeptide may be immobilized on a solid support and contacted with patient sera to allow binding of 5 antibodies within the sera to the immobilized polypeptide. Unbound sera may then be removed and bound antibodies detected using, for example, ¹²⁵I-labeled Protein A.

As noted above, a composition may comprise a variant of a native breast tumor protein. A polypeptide "variant," as used herein, is a polypeptide that differs from a native breast tumor protein in one or more substitutions, deletions, additions 10 and/or insertions, such that the immunogenicity of the polypeptide is not substantially diminished. In other words, the ability of a variant to react with antigen-specific antisera may be enhanced or unchanged, relative to the native protein, or may be diminished by less than 50%, and preferably less than 20%, relative to the native protein. Such variants may generally be identified by modifying one of the above 15 polypeptide sequences and evaluating the reactivity of the modified polypeptide with antigen-specific antibodies or antisera as described herein. Preferred variants include those in which one or more portions, such as an N-terminal leader sequence or transmembrane domain, have been removed. Other preferred variants include variants in which a small portion (e.g., 1-30 amino acids, preferably 5-15 amino acids) has been 20 removed from the N- and/or C-terminal of the mature protein.

Polypeptide variants encompassed by the present invention include those exhibiting at least about 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or more identity (determined as described above) to the polypeptides disclosed herein.

25 Preferably, a variant contains conservative substitutions. A "conservative substitution" is one in which an amino acid is substituted for another amino acid that has similar properties, such that one skilled in the art of peptide chemistry would expect the secondary structure and hydropathic nature of the polypeptide to be substantially unchanged. Amino acid substitutions may generally be 30 made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphipathic nature of the residues. For example, negatively

charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine and valine; glycine and alanine; asparagine and glutamine; and serine, threonine, phenylalanine and tyrosine.

5 Other groups of amino acids that may represent conservative changes include: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his; and (5) phe, tyr, trp, his. A variant may also, or alternatively, contain nonconservative changes. In a preferred embodiment, variant polypeptides differ from a native sequence by substitution, deletion or addition of five amino acids or fewer.

10 Variants may also (or alternatively) be modified by, for example, the deletion or addition of amino acids that have minimal influence on the immunogenicity, secondary structure and hydrophobic nature of the polypeptide.

As noted above, polypeptides may comprise a signal (or leader) sequence at the N-terminal end of the protein, which co-translationally or post-15 translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification or identification of the polypeptide (e.g., poly-His), or to enhance binding of the polypeptide to a solid support. For example, a polypeptide may be conjugated to an immunoglobulin Fc region.

Polypeptides may be prepared using any of a variety of well known 20 techniques. Recombinant polypeptides encoded by DNA sequences as described above may be readily prepared from the DNA sequences using any of a variety of expression vectors known to those of ordinary skill in the art. Expression may be achieved in any appropriate host cell that has been transformed or transfected with an expression vector containing a DNA molecule that encodes a recombinant polypeptide. Suitable host 25 cells include prokaryotes, yeast, and higher eukaryotic cells, such as mammalian cells and plant cells. Preferably, the host cells employed are *E. coli*, yeast or a mammalian cell line such as COS or CHO. Supernatants from suitable host/vector systems which secrete recombinant protein or polypeptide into culture media may be first concentrated using a commercially available filter. Following concentration, the concentrate may be 30 applied to a suitable purification matrix such as an affinity matrix or an ion exchange

resin. Finally, one or more reverse phase HPLC steps can be employed to further purify a recombinant polypeptide.

Portions and other variants having less than about 100 amino acids, and generally less than about 50 amino acids, may also be generated by synthetic means, 5 using techniques well known to those of ordinary skill in the art. For example, such polypeptides may be synthesized using any of the commercially available solid-phase techniques, such as the Merrifield solid-phase synthesis method, where amino acids are sequentially added to a growing amino acid chain. *See* Merrifield, *J. Am. Chem. Soc.* 85:2149-2146, 1963. Equipment for automated synthesis of polypeptides is 10 commercially available from suppliers such as Perkin Elmer/Applied BioSystems Division (Foster City, CA), and may be operated according to the manufacturer's instructions.

Within certain specific embodiments, a polypeptide may be a fusion protein that comprises multiple polypeptides as described herein, or that comprises at 15 least one polypeptide as described herein and an unrelated sequence, such as a known tumor protein. A fusion partner may, for example, assist in providing T helper epitopes (an immunological fusion partner), preferably T helper epitopes recognized by humans, or may assist in expressing the protein (an expression enhancer) at higher yields than the native recombinant protein. Certain preferred fusion partners are both 20 immunological and expression enhancing fusion partners. Other fusion partners may be selected so as to increase the solubility of the protein or to enable the protein to be targeted to desired intracellular compartments. Still further fusion partners include affinity tags, which facilitate purification of the protein.

Fusion proteins may generally be prepared using standard techniques, 25 including chemical conjugation. Preferably, a fusion protein is expressed as a recombinant protein, allowing the production of increased levels, relative to a non-fused protein, in an expression system. Briefly, DNA sequences encoding the polypeptide components may be assembled separately, and ligated into an appropriate expression vector. The 3' end of the DNA sequence encoding one polypeptide component is 30 ligated, with or without a peptide linker, to the 5' end of a DNA sequence encoding the second polypeptide component so that the reading frames of the sequences are in phase.

This permits translation into a single fusion protein that retains the biological activity of both component polypeptides.

A peptide linker sequence may be employed to separate the first and second polypeptide components by a distance sufficient to ensure that each polypeptide 5 folds into its secondary and tertiary structures. Such a peptide linker sequence is incorporated into the fusion protein using standard techniques well known in the art. Suitable peptide linker sequences may be chosen based on the following factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first and second 10 polypeptides; and (3) the lack of hydrophobic or charged residues that might react with the polypeptide functional epitopes. Preferred peptide linker sequences contain Gly, Asn and Ser residues. Other near neutral amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea et al., *Gene* 40:39-46, 1985; Murphy et al., 15 *Proc. Natl. Acad. Sci. USA* 83:8258-8262, 1986; U.S. Patent No. 4,935,233 and U.S. Patent No. 4,751,180. The linker sequence may generally be from 1 to about 50 amino acids in length. Linker sequences are not required when the first and second polypeptides have non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference.

20 The ligated DNA sequences are operably linked to suitable transcriptional or translational regulatory elements. The regulatory elements responsible for expression of DNA are located only 5' to the DNA sequence encoding the first polypeptides. Similarly, stop codons required to end translation and transcription termination signals are only present 3' to the DNA sequence encoding the 25 second polypeptide.

Fusion proteins are also provided. Such proteins comprise a polypeptide as described herein together with an unrelated immunogenic protein. Preferably the immunogenic protein is capable of eliciting a recall response. Examples of such proteins include tetanus, tuberculosis and hepatitis proteins (see, for example, Stoute 30 et al. *New Engl. J. Med.*, 336:86-91, 1997).

Within preferred embodiments, an immunological fusion partner is derived from protein D, a surface protein of the gram-negative bacterium *Haemophilus influenzae* B (WO 91/18926). Preferably, a protein D derivative comprises approximately the first third of the protein (e.g., the first N-terminal 100-110 amino acids), and a protein D derivative may be lipidated. Within certain preferred embodiments, the first 109 residues of a Lipoprotein D fusion partner is included on the N-terminus to provide the polypeptide with additional exogenous T-cell epitopes and to increase the expression level in *E. coli* (thus functioning as an expression enhancer). The lipid tail ensures optimal presentation of the antigen to antigen presenting cells.

10 Other fusion partners include the non-structural protein from influenzae virus, NS1 (hemagglutinin). Typically, the N-terminal 81 amino acids are used, although different fragments that include T-helper epitopes may be used.

In another embodiment, the immunological fusion partner is the protein known as LYTA, or a portion thereof (preferably a C-terminal portion). LYTA is derived from *Streptococcus pneumoniae*, which synthesizes an N-acetyl-L-alanine amidase known as amidase LYTA (encoded by the *LytA* gene; *Gene* 43:265-292, 1986). LYTA is an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal domain of the LYTA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE. This property has been exploited for the development of *E. coli* C-LYTA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LYTA fragment at the amino terminus has been described (see *Biotechnology* 10:795-798, 1992). Within a preferred embodiment, a repeat portion of LYTA may be incorporated into a fusion protein. A repeat portion is found in the C-terminal region starting at residue 178. A particularly preferred repeat portion incorporates residues 188-305.

In general, polypeptides (including fusion proteins) and polynucleotides as described herein are isolated. An "isolated" polypeptide or polynucleotide is one that is removed from its original environment. For example, a naturally-occurring protein is isolated if it is separated from some or all of the coexisting materials in the natural system. Preferably, such polypeptides are at least about 90% pure, more preferably at

least about 95% pure and most preferably at least about 99% pure. A polynucleotide is considered to be isolated if, for example, it is cloned into a vector that is not a part of the natural environment.

BINDING AGENTS

5 The present invention further provides agents, such as antibodies and antigen-binding fragments thereof, that specifically bind to a breast tumor protein. As used herein, an antibody, or antigen-binding fragment thereof, is said to "specifically bind" to a breast tumor protein if it reacts at a detectable level (within, for example, an ELISA) with a breast tumor protein, and does not react detectably with unrelated 10 proteins under similar conditions. As used herein, "binding" refers to a noncovalent association between two separate molecules such that a complex is formed. The ability to bind may be evaluated by, for example, determining a binding constant for the formation of the complex. The binding constant is the value obtained when the concentration of the complex is divided by the product of the component 15 concentrations. In general, two compounds are said to "bind," in the context of the present invention, when the binding constant for complex formation exceeds about 10^3 L/mol. The binding constant may be determined using methods well known in the art.

Binding agents may be further capable of differentiating between patients with and without a cancer, such as breast cancer, using the representative assays 20 provided herein. In other words, antibodies or other binding agents that bind to a breast tumor protein will generate a signal indicating the presence of a cancer in at least about 20% of patients with the disease, and will generate a negative signal indicating the absence of the disease in at least about 90% of individuals without the cancer. To determine whether a binding agent satisfies this requirement, biological samples (e.g., 25 blood, sera, sputum, urine and/or tumor biopsies) from patients with and without a cancer (as determined using standard clinical tests) may be assayed as described herein for the presence of polypeptides that bind to the binding agent. It will be apparent that a statistically significant number of samples with and without the disease should be assayed. Each binding agent should satisfy the above criteria; however, those of

ordinary skill in the art will recognize that binding agents may be used in combination to improve sensitivity.

Any agent that satisfies the above requirements may be a binding agent. For example, a binding agent may be a ribosome, with or without a peptide component, 5 an RNA molecule or a polypeptide. In a preferred embodiment, a binding agent is an antibody or an antigen-binding fragment thereof. Antibodies may be prepared by any of a variety of techniques known to those of ordinary skill in the art. *See, e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, antibodies can be produced by cell culture techniques, including the generation 10 of monoclonal antibodies as described herein, or via transfection of antibody genes into suitable bacterial or mammalian cell hosts, in order to allow for the production of recombinant antibodies. In one technique, an immunogen comprising the polypeptide is initially injected into any of a wide variety of mammals (*e.g.*, mice, rats, rabbits, sheep or goats). In this step, the polypeptides of this invention may serve as the immunogen 15 without modification. Alternatively, particularly for relatively short polypeptides, a superior immune response may be elicited if the polypeptide is joined to a carrier protein, such as bovine serum albumin or keyhole limpet hemocyanin. The immunogen is injected into the animal host, preferably according to a predetermined schedule incorporating one or more booster immunizations, and the animals are bled periodically. 20 Polyclonal antibodies specific for the polypeptide may then be purified from such antisera by, for example, affinity chromatography using the polypeptide coupled to a suitable solid support.

Monoclonal antibodies specific for an antigenic polypeptide of interest may be prepared, for example, using the technique of Kohler and Milstein, *Eur. J. 25 Immunol.* 6:511-519, 1976, and improvements thereto. Briefly, these methods involve the preparation of immortal cell lines capable of producing antibodies having the desired specificity (*i.e.*, reactivity with the polypeptide of interest). Such cell lines may be produced, for example, from spleen cells obtained from an animal immunized as described above. The spleen cells are then immortalized by, for example, fusion with a 30 myeloma cell fusion partner, preferably one that is syngeneic with the immunized animal. A variety of fusion techniques may be employed. For example, the spleen cells

and myeloma cells may be combined with a nonionic detergent for a few minutes and then plated at low density on a selective medium that supports the growth of hybrid cells, but not myeloma cells. A preferred selection technique uses HAT (hypoxanthine, aminopterin, thymidine) selection. After a sufficient time, usually about 1 to 2 weeks, 5 colonies of hybrids are observed. Single colonies are selected and their culture supernatants tested for binding activity against the polypeptide. Hybridomas having high reactivity and specificity are preferred.

Monoclonal antibodies may be isolated from the supernatants of growing hybridoma colonies. In addition, various techniques may be employed to enhance the 10 yield, such as injection of the hybridoma cell line into the peritoneal cavity of a suitable vertebrate host, such as a mouse. Monoclonal antibodies may then be harvested from the ascites fluid or the blood. Contaminants may be removed from the antibodies by conventional techniques, such as chromatography, gel filtration, precipitation, and extraction. The polypeptides of this invention may be used in the purification process 15 in, for example, an affinity chromatography step.

Within certain embodiments, the use of antigen-binding fragments of antibodies may be preferred. Such fragments include Fab fragments, which may be prepared using standard techniques. Briefly, immunoglobulins may be purified from rabbit serum by affinity chromatography on Protein A bead columns (Harlow and Lane, 20 *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988) and digested by papain to yield Fab and Fc fragments. The Fab and Fc fragments may be separated by affinity chromatography on protein A bead columns.

Monoclonal antibodies of the present invention may be coupled to one or more therapeutic agents. Suitable agents in this regard include radionuclides, 25 differentiation inducers, drugs, toxins, and derivatives thereof. Preferred radionuclides include ⁹⁰Y, ¹²³I, ¹²⁵I, ¹³¹I, ¹⁸⁶Re, ¹⁸⁸Re, ²¹¹At, and ²¹²Bi. Preferred drugs include methotrexate, and pyrimidine and purine analogs. Preferred differentiation inducers include phorbol esters and butyric acid. Preferred toxins include ricin, abrin, diphtheria toxin, cholera toxin, gelonin, *Pseudomonas* exotoxin, *Shigella* toxin, and pokeweed 30 antiviral protein.

A therapeutic agent may be coupled (*e.g.*, covalently bonded) to a suitable monoclonal antibody either directly or indirectly (*e.g.*, via a linker group). A direct reaction between an agent and an antibody is possible when each possesses a substituent capable of reacting with the other. For example, a nucleophilic group, such 5 as an amino or sulfhydryl group, on one may be capable of reacting with a carbonyl-containing group, such as an anhydride or an acid halide, or with an alkyl group containing a good leaving group (*e.g.*, a halide) on the other.

Alternatively, it may be desirable to couple a therapeutic agent and an antibody via a linker group. A linker group can function as a spacer to distance an 10 antibody from an agent in order to avoid interference with binding capabilities. A linker group can also serve to increase the chemical reactivity of a substituent on an agent or an antibody, and thus increase the coupling efficiency. An increase in chemical reactivity may also facilitate the use of agents, or functional groups on agents, which otherwise would not be possible.

15 It will be evident to those skilled in the art that a variety of bifunctional or polyfunctional reagents, both homo- and hetero-functional (such as those described in the catalog of the Pierce Chemical Co., Rockford, IL), may be employed as the linker group. Coupling may be effected, for example, through amino groups, carboxyl groups, sulfhydryl groups or oxidized carbohydrate residues. There are numerous references 20 describing such methodology, *e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.

Where a therapeutic agent is more potent when free from the antibody portion of the immunoconjugates of the present invention, it may be desirable to use a linker group which is cleavable during or upon internalization into a cell. A number of different cleavable linker groups have been described. The mechanisms for the 25 intracellular release of an agent from these linker groups include cleavage by reduction of a disulfide bond (*e.g.*, U.S. Patent No. 4,489,710, to Spitzer), by irradiation of a photolabile bond (*e.g.*, U.S. Patent No. 4,625,014, to Senter et al.), by hydrolysis of derivatized amino acid side chains (*e.g.*, U.S. Patent No. 4,638,045, to Kohn et al.), by serum complement-mediated hydrolysis (*e.g.*, U.S. Patent No. 4,671,958, to Rodwell 30 et al.), and acid-catalyzed hydrolysis (*e.g.*, U.S. Patent No. 4,569,789, to Blattler et al.).

It may be desirable to couple more than one agent to an antibody. In one embodiment, multiple molecules of an agent are coupled to one antibody molecule. In another embodiment, more than one type of agent may be coupled to one antibody. Regardless of the particular embodiment, immunoconjugates with more than one agent 5 may be prepared in a variety of ways. For example, more than one agent may be coupled directly to an antibody molecule, or linkers that provide multiple sites for attachment can be used. Alternatively, a carrier can be used.

A carrier may bear the agents in a variety of ways, including covalent bonding either directly or via a linker group. Suitable carriers include proteins such as 10 albumins (e.g., U.S. Patent No. 4,507,234, to Kato et al.), peptides and polysaccharides such as aminodextran (e.g., U.S. Patent No. 4,699,784, to Shih et al.). A carrier may also bear an agent by noncovalent bonding or by encapsulation, such as within a liposome vesicle (e.g., U.S. Patent Nos. 4,429,008 and 4,873,088). Carriers specific for 15 radionuclide agents include radiohalogenated small molecules and chelating compounds. For example, U.S. Patent No. 4,735,792 discloses representative radiohalogenated small molecules and their synthesis. A radionuclide chelate may be formed from chelating compounds that include those containing nitrogen and sulfur atoms as the donor atoms for binding the metal, or metal oxide, radionuclide. For 20 example, U.S. Patent No. 4,673,562, to Davison et al. discloses representative chelating compounds and their synthesis.

A variety of routes of administration for the antibodies and immunoconjugates may be used. Typically, administration will be intravenous, intramuscular, subcutaneous or in the bed of a resected tumor. It will be evident that the precise dose of the antibody/immunoconjugate will vary depending upon the antibody 25 used, the antigen density on the tumor, and the rate of clearance of the antibody.

T CELLS

Immunotherapeutic compositions may also, or alternatively, comprise T cells specific for a breast tumor protein. Such cells may generally be prepared *in vitro* or *ex vivo*, using standard procedures. For example, T cells may be isolated from bone 30 marrow, peripheral blood, or a fraction of bone marrow or peripheral blood of a patient,

using a commercially available cell separation system, such as the Isolex™ System, available from Nexell Therapeutics, Inc. (Irvine, CA; see also U.S. Patent No. 5,240,856; U.S. Patent No. 5,215,926; WO 89/06280; WO 91/16116 and WO 92/07243). Alternatively, T cells may be derived from related or unrelated humans, 5 non-human mammals, cell lines or cultures.

T cells may be stimulated with a breast tumor polypeptide, polynucleotide encoding a breast tumor polypeptide and/or an antigen presenting cell (APC) that expresses such a polypeptide. Such stimulation is performed under conditions and for a time sufficient to permit the generation of T cells that are specific 10 for the polypeptide. Preferably, a breast tumor polypeptide or polynucleotide is present within a delivery vehicle, such as a microsphere, to facilitate the generation of specific T cells.

T cells are considered to be specific for a breast tumor polypeptide if the T cells specifically proliferate, secrete cytokines or kill target cells coated with the 15 polypeptide or expressing a gene encoding the polypeptide. T cell specificity may be evaluated using any of a variety of standard techniques. For example, within a chromium release assay or proliferation assay, a stimulation index of more than two fold increase in lysis and/or proliferation, compared to negative controls, indicates T cell specificity. Such assays may be performed, for example, as described in Chen et 20 al., *Cancer Res.* 54:1065-1070, 1994. Alternatively, detection of the proliferation of T cells may be accomplished by a variety of known techniques. For example, T cell proliferation can be detected by measuring an increased rate of DNA synthesis (e.g., by pulse-labeling cultures of T cells with tritiated thymidine and measuring the amount of tritiated thymidine incorporated into DNA). Contact with a breast tumor polypeptide 25 (100 ng/ml - 100 µg/ml, preferably 200 ng/ml - 25 µg/ml) for 3 - 7 days should result in at least a two fold increase in proliferation of the T cells. Contact as described above for 2-3 hours should result in activation of the T cells, as measured using standard cytokine assays in which a two fold increase in the level of cytokine release (e.g., TNF or IFN- γ) is indicative of T cell activation (see Coligan et al., *Current Protocols in 30 Immunology*, vol. 1, Wiley Interscience (Greene 1998)). T cells that have been activated in response to a breast tumor polypeptide, polynucleotide or polypeptide-

expressing APC may be CD4⁺ and/or CD8⁺. Breast tumor protein-specific T cells may be expanded using standard techniques. Within preferred embodiments, the T cells are derived from a patient, a related donor or an unrelated donor, and are administered to the patient following stimulation and expansion.

5 For therapeutic purposes, CD4⁺ or CD8⁺ T cells that proliferate in response to a breast tumor polypeptide, polynucleotide or APC can be expanded in number either *in vitro* or *in vivo*. Proliferation of such T cells *in vitro* may be accomplished in a variety of ways. For example, the T cells can be re-exposed to a breast tumor polypeptide, or a short peptide corresponding to an immunogenic portion

10 of such a polypeptide, with or without the addition of T cell growth factors, such as interleukin-2, and/or stimulator cells that synthesize a breast tumor polypeptide. Alternatively, one or more T cells that proliferate in the presence of a breast tumor protein can be expanded in number by cloning. Methods for cloning cells are well known in the art, and include limiting dilution.

15 PHARMACEUTICAL COMPOSITIONS

In additional embodiments, the present invention concerns formulation of one or more of the polynucleotide, polypeptide, T-cell and/or antibody compositions disclosed herein in pharmaceutically-acceptable solutions for administration to a cell or an animal, either alone, or in combination with one or more other modalities of therapy.

20 It will also be understood that, if desired, the nucleic acid segment, RNA, DNA or PNA compositions that express a polypeptide as disclosed herein may be administered in combination with other agents as well, such as, *e.g.*, other proteins or polypeptides or various pharmaceutically-active agents. In fact, there is virtually no limit to other components that may also be included, given that the additional agents do

25 not cause a significant adverse effect upon contact with the target cells or host tissues. The compositions may thus be delivered along with various other agents as required in the particular instance. Such compositions may be purified from host cells or other biological sources, or alternatively may be chemically synthesized as described herein. Likewise, such compositions may further comprise substituted or derivatized RNA or

30 DNA compositions.

Formulation of pharmaceutically-acceptable excipients and carrier solutions is well-known to those of skill in the art, as is the development of suitable dosing and treatment regimens for using the particular compositions described herein in a variety of treatment regimens, including *e.g.*, oral, parenteral, intravenous, intranasal, 5 and intramuscular administration and formulation.

1. ORAL DELIVERY

In certain applications, the pharmaceutical compositions disclosed herein may be delivered *via* oral administration to an animal. As such, these compositions may be formulated with an inert diluent or with an assimilable edible carrier, or they 10 may be enclosed in hard- or soft-shell gelatin capsule, or they may be compressed into tablets, or they may be incorporated directly with the food of the diet.

The active compounds may even be incorporated with excipients and used in the form of ingestible tablets, buccal tables, troches, capsules, elixirs, suspensions, syrups, wafers, and the like (Mathiowitz *et al.*, 1997; Hwang *et al.*, 1998; 15 U. S. Patent 5,641,515; U. S. Patent 5,580,579 and U. S. Patent 5,792,451, each specifically incorporated herein by reference in its entirety). The tablets, troches, pills, capsules and the like may also contain the following: a binder, as gum tragacanth, acacia, cornstarch, or gelatin; excipients, such as dicalcium phosphate; a disintegrating agent, such as corn starch, potato starch, alginic acid and the like; a lubricant, such as 20 magnesium stearate; and a sweetening agent, such as sucrose, lactose or saccharin may be added or a flavoring agent, such as peppermint, oil of wintergreen, or cherry flavoring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, 25 tablets, pills, or capsules may be coated with shellac, sugar, or both. A syrup of elixir may contain the active compound sucrose as a sweetening agent methyl and propylparabens as preservatives, a dye and flavoring, such as cherry or orange flavor. Of course, any material used in preparing any dosage unit form should be pharmaceutically pure and substantially non-toxic in the amounts employed. In

addition, the active compounds may be incorporated into sustained-release preparation and formulations.

Typically, these formulations may contain at least about 0.1% of the active compound or more, although the percentage of the active ingredient(s) may, of course, be varied and may conveniently be between about 1 or 2% and about 60% or 70% or more of the weight or volume of the total formulation. Naturally, the amount of active compound(s) in each therapeutically useful composition may be prepared in such a way that a suitable dosage will be obtained in any given unit dose of the compound. Factors such as solubility, bioavailability, biological half-life, route of administration, product shelf life, as well as other pharmacological considerations will be contemplated by one skilled in the art of preparing such pharmaceutical formulations, and as such, a variety of dosages and treatment regimens may be desirable.

For oral administration the compositions of the present invention may alternatively be incorporated with one or more excipients in the form of a mouthwash, dentifrice, buccal tablet, oral spray, or sublingual orally-administered formulation. For example, a mouthwash may be prepared incorporating the active ingredient in the required amount in an appropriate solvent, such as a sodium borate solution (Dobell's Solution). Alternatively, the active ingredient may be incorporated into an oral solution such as one containing sodium borate, glycerin and potassium bicarbonate, or dispersed in a dentifrice, or added in a therapeutically-effective amount to a composition that may include water, binders, abrasives, flavoring agents, foaming agents, and humectants. Alternatively the compositions may be fashioned into a tablet or solution form that may be placed under the tongue or otherwise dissolved in the mouth.

2. INJECTABLE DELIVERY

In certain circumstances it will be desirable to deliver the pharmaceutical compositions disclosed herein parenterally, intravenously, intramuscularly, or even intraperitoneally as described in U. S. Patent 5,543,158; U. S. Patent 5,641,515 and U. S. Patent 5,399,363 (each specifically incorporated herein by reference in its entirety). Solutions of the active compounds as free base or pharmacologically acceptable salts may be prepared in water suitably mixed with a surfactant, such as

hydroxypropylcellulose. Dispersions may also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations contain a preservative to prevent the growth of microorganisms.

5 The pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions (U. S. Patent 5,466,468, specifically incorporated herein by reference in its entirety). In all cases the form must be sterile and must be fluid to the extent that easy syringability exists. It must be stable
10 under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (e.g., glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and/or vegetable oils. Proper fluidity may be maintained, for
15 example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. The prevention of the action of microorganisms can be facilitated by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars
20 or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

For parenteral administration in an aqueous solution, for example, the solution should be suitably buffered if necessary and the liquid diluent first rendered
25 isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In this connection, a sterile aqueous medium that can be employed will be known to those of skill in the art in light of the present disclosure. For example, one dosage may be dissolved in 1 ml of isotonic NaCl solution and either added to 1000 ml
30 of hypodermoclysis fluid or injected at the proposed site of infusion, (see for example, "Remington's Pharmaceutical Sciences" 15th Edition, pages 1035-1038 and 1570-

1580). Some variation in dosage will necessarily occur depending on the condition of the subject being treated. The person responsible for administration will, in any event, determine the appropriate dose for the individual subject. Moreover, for human administration, preparations should meet sterility, pyrogenicity, and the general safety 5 and purity standards as required by FDA Office of Biologics standards.

Sterile injectable solutions are prepared by incorporating the active compounds in the required amount in the appropriate solvent with various of the other ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the various sterilized active ingredients into a 10 sterile vehicle which contains the basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum-drying and freeze-drying techniques which yield a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered 15 solution thereof.

The compositions disclosed herein may be formulated in a neutral or salt form. Pharmaceutically-acceptable salts, include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, 20 oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, histidine, procaine and the like. Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount 25 as is therapeutically effective. The formulations are easily administered in a variety of dosage forms such as injectable solutions, drug-release capsules, and the like.

As used herein, "carrier" includes any and all solvents, dispersion media, vehicles, coatings, diluents, antibacterial and antifungal agents, isotonic and absorption delaying agents, buffers, carrier solutions, suspensions, colloids, and the like. The use 30 of such media and agents for pharmaceutical active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active

ingredient, its use in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions.

The phrase "pharmaceutically-acceptable" refers to molecular entities and compositions that do not produce an allergic or similar untoward reaction when 5 administered to a human. The preparation of an aqueous composition that contains a protein as an active ingredient is well understood in the art. Typically, such compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid prior to injection can also be prepared. The preparation can also be emulsified.

10 3. NASAL DELIVERY

In certain embodiments, the pharmaceutical compositions may be delivered by intranasal sprays, inhalation, and/or other aerosol delivery vehicles. Methods for delivering genes, nucleic acids, and peptide compositions directly to the lungs *via* nasal aerosol sprays has been described *e.g.*, in U. S. Patent 5,756,353 and U. 15 S. Patent 5,804,212 (each specifically incorporated herein by reference in its entirety). Likewise, the delivery of drugs using intranasal microparticle resins (Takenaga *et al.*, 1998) and lysophosphatidyl-glycerol compounds (U. S. Patent 5,725,871, specifically incorporated herein by reference in its entirety) are also well-known in the pharmaceutical arts. Likewise, transmucosal drug delivery in the form of a 20 polytetrafluoroethylene support matrix is described in U. S. Patent 5,780,045 (specifically incorporated herein by reference in its entirety).

4. LIPOSOME-, NANOCAPSULE-, AND MICROPARTICLE-MEDIATED DELIVERY

In certain embodiments, the inventors contemplate the use of liposomes, nanocapsules, microparticles, microspheres, lipid particles, vesicles, and the like, for the 25 introduction of the compositions of the present invention into suitable host cells. In particular, the compositions of the present invention may be formulated for delivery either encapsulated in a lipid particle, a liposome, a vesicle, a nanosphere, or a nanoparticle or the like.

Such formulations may be preferred for the introduction of pharmaceutically-acceptable formulations of the nucleic acids or constructs disclosed herein. The formation and use of liposomes is generally known to those of skill in the art (see for example, Couvreur *et al.*, 1977; Couvreur, 1988; Lasic, 1998; which 5 describes the use of liposomes and nanocapsules in the targeted antibiotic therapy for intracellular bacterial infections and diseases). Recently, liposomes were developed with improved serum stability and circulation half-times (Gabizon and Papahadjopoulos, 1988; Allen and Choun, 1987; U. S. Patent 5,741,516, specifically incorporated herein by reference in its entirety). Further, various methods of liposome 10 and liposome like preparations as potential drug carriers have been reviewed (Takakura, 1998; Chandran *et al.*, 1997; Margalit, 1995; U. S. Patent 5,567,434; U. S. Patent 5,552,157; U. S. Patent 5,565,213; U. S. Patent 5,738,868 and U. S. Patent 5,795,587, each specifically incorporated herein by reference in its entirety).

Liposomes have been used successfully with a number of cell types that 15 are normally resistant to transfection by other procedures including T cell suspensions, primary hepatocyte cultures and PC 12 cells (Renneisen *et al.*, 1990; Muller *et al.*, 1990). In addition, liposomes are free of the DNA length constraints that are typical of viral-based delivery systems. Liposomes have been used effectively to introduce genes, 20 drugs (Heath and Martin, 1986; Heath *et al.*, 1986; Balazsorvits *et al.*, 1989; Fresta and Puglisi, 1996), radiotherapeutic agents (Pikul *et al.*, 1987), enzymes (Imaizumi *et al.*, 1990a; Imaizumi *et al.*, 1990b), viruses (Faller and Baltimore, 1984), transcription factors and allosteric effectors (Nicolau and Gersonde, 1979) into a variety of cultured 25 cell lines and animals. In addition, several successful clinical trials examining the effectiveness of liposome-mediated drug delivery have been completed (Lopez-Berestein *et al.*, 1985a; 1985b; Coune, 1988; Sculier *et al.*, 1988). Furthermore, several studies suggest that the use of liposomes is not associated with autoimmune responses, toxicity or gonadal localization after systemic delivery (Mori and Fukatsu, 1992).

Liposomes are formed from phospholipids that are dispersed in an aqueous medium and spontaneously form multilamellar concentric bilayer vesicles (also 30 termed multilamellar vesicles (MLVs). MLVs generally have diameters of from 25 nm to 4 μ m. Sonication of MLVs results in the formation of small unilamellar vesicles

(SUVs) with diameters in the range of 200 to 500 Å, containing an aqueous solution in the core.

Liposomes bear resemblance to cellular membranes and are contemplated for use in connection with the present invention as carriers for the peptide 5 compositions. They are widely suitable as both water- and lipid-soluble substances can be entrapped, *i.e.* in the aqueous spaces and within the bilayer itself, respectively. It is possible that the drug-bearing liposomes may even be employed for site-specific delivery of active agents by selectively modifying the liposomal formulation.

In addition to the teachings of Couvreur *et al.* (1977; 1988), the 10 following information may be utilized in generating liposomal formulations. Phospholipids can form a variety of structures other than liposomes when dispersed in water, depending on the molar ratio of lipid to water. At low ratios the liposome is the preferred structure. The physical characteristics of liposomes depend on pH, ionic strength and the presence of divalent cations. Liposomes can show low permeability to 15 ionic and polar substances, but at elevated temperatures undergo a phase transition which markedly alters their permeability. The phase transition involves a change from a closely packed, ordered structure, known as the gel state, to a loosely packed, less-ordered structure, known as the fluid state. This occurs at a characteristic phase-transition temperature and results in an increase in permeability to ions, sugars and 20 drugs.

In addition to temperature, exposure to proteins can alter the permeability of liposomes. Certain soluble proteins, such as cytochrome c, bind, deform and penetrate the bilayer, thereby causing changes in permeability. Cholesterol inhibits this penetration of proteins, apparently by packing the phospholipids more 25 tightly. It is contemplated that the most useful liposome formations for antibiotic and inhibitor delivery will contain cholesterol.

The ability to trap solutes varies between different types of liposomes. For example, MLVs are moderately efficient at trapping solutes, but SUVs are extremely inefficient. SUVs offer the advantage of homogeneity and reproducibility in 30 size distribution, however, and a compromise between size and trapping efficiency is

offered by large unilamellar vesicles (LUVs). These are prepared by ether evaporation and are three to four times more efficient at solute entrapment than MLVs.

In addition to liposome characteristics, an important determinant in entrapping compounds is the physicochemical properties of the compound itself. Polar 5 compounds are trapped in the aqueous spaces and nonpolar compounds bind to the lipid bilayer of the vesicle. Polar compounds are released through permeation or when the bilayer is broken, but nonpolar compounds remain affiliated with the bilayer unless it is disrupted by temperature or exposure to lipoproteins. Both types show maximum efflux rates at the phase transition temperature.

10 Liposomes interact with cells *via* four different mechanisms: endocytosis by phagocytic cells of the reticuloendothelial system such as macrophages and neutrophils; adsorption to the cell surface, either by nonspecific weak hydrophobic or electrostatic forces, or by specific interactions with cell-surface components; fusion with the plasma cell membrane by insertion of the lipid bilayer of the liposome into the 15 plasma membrane, with simultaneous release of liposomal contents into the cytoplasm; and by transfer of liposomal lipids to cellular or subcellular membranes, or vice versa, without any association of the liposome contents. It often is difficult to determine which mechanism is operative and more than one may operate at the same time.

20 The fate and disposition of intravenously injected liposomes depend on their physical properties, such as size, fluidity, and surface charge. They may persist in tissues for h or days, depending on their composition, and half lives in the blood range from min to several h. Larger liposomes, such as MLVs and LUVs, are taken up rapidly by phagocytic cells of the reticuloendothelial system, but physiology of the circulatory system restrains the exit of such large species at most sites. They can exit 25 only in places where large openings or pores exist in the capillary endothelium, such as the sinusoids of the liver or spleen. Thus, these organs are the predominate site of uptake. On the other hand, SUVs show a broader tissue distribution but still are sequestered highly in the liver and spleen. In general, this *in vivo* behavior limits the potential targeting of liposomes to only those organs and tissues accessible to their large 30 size. These include the blood, liver, spleen, bone marrow, and lymphoid organs.

Targeting is generally not a limitation in terms of the present invention. However, should specific targeting be desired, methods are available for this to be accomplished. Antibodies may be used to bind to the liposome surface and to direct the antibody and its drug contents to specific antigenic receptors located on a particular 5 cell-type surface. Carbohydrate determinants (glycoprotein or glycolipid cell-surface components that play a role in cell-cell recognition, interaction and adhesion) may also be used as recognition sites as they have potential in directing liposomes to particular cell types. Mostly, it is contemplated that intravenous injection of liposomal preparations would be used, but other routes of administration are also conceivable.

10 Alternatively, the invention provides for pharmaceutically-acceptable nanocapsule formulations of the compositions of the present invention. Nanocapsules can generally entrap compounds in a stable and reproducible way (Henry-Michelland *et al.*, 1987; Quintanar-Guerrero *et al.*, 1998; Douglas *et al.*, 1987). To avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 15 0.1 μm) should be designed using polymers able to be degraded *in vivo*. Biodegradable polyalkyl-cyanoacrylate nanoparticles that meet these requirements are contemplated for use in the present invention. Such particles may be easily made, as described (Couvreur *et al.*, 1980; 1988; zur Muhlen *et al.*, 1998; Zambaux *et al.* 1998; Pinto-Alphandry *et al.*, 1995 and U. S. Patent 5,145,684, specifically incorporated herein by 20 reference in its entirety).

VACCINES

In certain preferred embodiments of the present invention, vaccines are provided. The vaccines will generally comprise one or more pharmaceutical compositions, such as those discussed above, in combination with an immunostimulant. 25 An immunostimulant may be any substance that enhances or potentiates an immune response (antibody and/or cell-mediated) to an exogenous antigen. Examples of immunostimulants include adjuvants, biodegradable microspheres (e.g., polylactic galactide) and liposomes (into which the compound is incorporated; *see e.g.*, Fullerton, U.S. Patent No. 4,235,877). Vaccine preparation is generally described in, for example, 30 M.F. Powell and M.J. Newman, eds., "Vaccine Design (the subunit and adjuvant

approach)," Plenum Press (NY, 1995). Pharmaceutical compositions and vaccines within the scope of the present invention may also contain other compounds, which may be biologically active or inactive. For example, one or more immunogenic portions of other tumor antigens may be present, either incorporated into a fusion 5 polypeptide or as a separate compound, within the composition or vaccine.

Illustrative vaccines may contain DNA encoding one or more of the polypeptides as described above, such that the polypeptide is generated *in situ*. As noted above, the DNA may be present within any of a variety of delivery systems known to those of ordinary skill in the art, including nucleic acid expression systems, 10 bacteria and viral expression systems. Numerous gene delivery techniques are well known in the art, such as those described by Rolland, *Crit. Rev. Therap. Drug Carrier Systems* 15:143-198, 1998, and references cited therein. Appropriate nucleic acid expression systems contain the necessary DNA sequences for expression in the patient (such as a suitable promoter and terminating signal). Bacterial delivery systems involve 15 the administration of a bacterium (such as *Bacillus-Calmette-Guerrin*) that expresses an immunogenic portion of the polypeptide on its cell surface or secretes such an epitope. In a preferred embodiment, the DNA may be introduced using a viral expression system (e.g., vaccinia or other pox virus, retrovirus, or adenovirus), which may involve the use of a non-pathogenic (defective), replication competent virus. Suitable systems are 20 disclosed, for example, in Fisher-Hoch et al., *Proc. Natl. Acad. Sci. USA* 86:317-321, 1989; Flexner et al., *Ann. N.Y. Acad. Sci.* 569:86-103, 1989; Flexner et al., *Vaccine* 8:17-21, 1990; U.S. Patent Nos. 4,603,112, 4,769,330, and 5,017,487; WO 89/01973; U.S. Patent No. 4,777,127; GB 2,200,651; EP 0,345,242; WO 91/02805; Berkner, 25 *Biotechniques* 6:616-627, 1988; Rosenfeld et al., *Science* 252:431-434, 1991; Kolls et al., *Proc. Natl. Acad. Sci. USA* 91:215-219, 1994; Kass-Eisler et al., *Proc. Natl. Acad. Sci. USA* 90:11498-11502, 1993; Guzman et al., *Circulation* 88:2838-2848, 1993; and Guzman et al., *Cir. Res.* 73:1202-1207, 1993. Techniques for incorporating DNA into such expression systems are well known to those of ordinary skill in the art. The DNA may also be "naked," as described, for example, in Ulmer et al., *Science* 259:1745-1749, 30 1993 and reviewed by Cohen, *Science* 259:1691-1692, 1993. The uptake of naked DNA may be increased by coating the DNA onto biodegradable beads, which are

efficiently transported into the cells. It will be apparent that a vaccine may comprise both a polynucleotide and a polypeptide component. Such vaccines may provide for an enhanced immune response.

It will be apparent that a vaccine may contain pharmaceutically acceptable salts of the polynucleotides and polypeptides provided herein. Such salts may be prepared from pharmaceutically acceptable non-toxic bases, including organic bases (e.g., salts of primary, secondary and tertiary amines and basic amino acids) and inorganic bases (e.g., sodium, potassium, lithium, ammonium, calcium and magnesium salts).

10 While any suitable carrier known to those of ordinary skill in the art may be employed in the vaccine compositions of this invention, the type of carrier will vary depending on the mode of administration. Compositions of the present invention may be formulated for any appropriate manner of administration, including for example, topical, oral, nasal, intravenous, intracranial, intraperitoneal, subcutaneous or 15 intramuscular administration. For parenteral administration, such as subcutaneous injection, the carrier preferably comprises water, saline, alcohol, a fat, a wax or a buffer. For oral administration, any of the above carriers or a solid carrier, such as mannitol, lactose, starch, magnesium stearate, sodium saccharine, talcum, cellulose, glucose, sucrose, and magnesium carbonate, may be employed. Biodegradable microspheres 20 (e.g., polylactate polyglycolate) may also be employed as carriers for the pharmaceutical compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268; 5,075,109; 5,928,647; 5,811,128; 5,820,883; 5,853,763; 5,814,344 and 5,942,252. One may also employ a carrier comprising the particulate-protein complexes described in U.S. Patent No. 25 5,928,647, which are capable of inducing a class I-restricted cytotoxic T lymphocyte responses in a host.

Such compositions may also comprise buffers (e.g., neutral buffered saline or phosphate buffered saline), carbohydrates (e.g., glucose, mannose, sucrose or dextrans), mannitol, proteins, polypeptides or amino acids such as glycine, antioxidants, 30 bacteriostats, chelating agents such as EDTA or glutathione, adjuvants (e.g., aluminum hydroxide), solutes that render the formulation isotonic, hypotonic or weakly hypertonic

with the blood of a recipient, suspending agents, thickening agents and/or preservatives. Alternatively, compositions of the present invention may be formulated as a lyophilizate. Compounds may also be encapsulated within liposomes using well known technology.

5 Any of a variety of immunostimulants may be employed in the vaccines of this invention. For example, an adjuvant may be included. Most adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bordetella pertussis* or *Mycobacterium tuberculosis* derived proteins. Suitable 10 adjuvants are commercially available as, for example, Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; 15 cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF or interleukin-2, -7, or -12, may also be used as adjuvants.

Within the vaccines provided herein, the adjuvant composition is preferably designed to induce an immune response predominantly of the Th1 type. 20 High levels of Th1-type cytokines (e.g., IFN- γ , TNF α , IL-2 and IL-12) tend to favor the induction of cell mediated immune responses to an administered antigen. In contrast, high levels of Th2-type cytokines (e.g., IL-4, IL-5, IL-6 and IL-10) tend to favor the induction of humoral immune responses. Following application of a vaccine as provided herein, a patient will support an immune response that includes Th1- and Th2- 25 type responses. Within a preferred embodiment, in which a response is predominantly Th1-type, the level of Th1-type cytokines will increase to a greater extent than the level of Th2-type cytokines. The levels of these cytokines may be readily assessed using standard assays. For a review of the families of cytokines, see Mosmann and Coffman, *Ann. Rev. Immunol.* 7:145-173, 1989.

30 Preferred adjuvants for use in eliciting a predominantly Th1-type response include, for example, a combination of monophosphoryl lipid A, preferably 3-

de-O-acylated monophosphoryl lipid A (3D-MPL), together with an aluminum salt. MPL adjuvants are available from Corixa Corporation (Seattle, WA; *see* US Patent Nos. 4,436,727; 4,877,611; 4,866,034 and 4,912,094). CpG-containing oligonucleotides (in which the CpG dinucleotide is unmethylated) also induce a predominantly Th1 response. Such oligonucleotides are well known and are described, for example, in WO 96/02555, WO 99/33488 and U.S. Patent Nos. 6,008,200 and 5,856,462. Immunostimulatory DNA sequences are also described, for example, by Sato et al., *Science* 273:352, 1996. Another preferred adjuvant is a saponin, preferably QS21 (Aquila Biopharmaceuticals Inc., Framingham, MA), which may be used alone or in combination with other adjuvants. For example, an enhanced system involves the combination of a monophosphoryl lipid A and saponin derivative, such as the combination of QS21 and 3D-MPL as described in WO 94/00153, or a less reactogenic composition where the QS21 is quenched with cholesterol, as described in WO 96/33739. Other preferred formulations comprise an oil-in-water emulsion and tocopherol. A particularly potent adjuvant formulation involving QS21, 3D-MPL and tocopherol in an oil-in-water emulsion is described in WO 95/17210.

Other preferred adjuvants include Montanide ISA 720 (Seppic, France), SAF (Chiron, California, United States), ISCOMS (CSL), MF-59 (Chiron), the SBAS series of adjuvants (*e.g.*, SBAS-2 or SBAS-4, available from SmithKline Beecham, Rixensart, Belgium), Detox (Corixa, Hamilton, MT), RC-529 (Corixa, Hamilton, MT) and other aminoalkyl glucosaminide 4-phosphates (AGPs), such as those described in pending U.S. Patent Application Serial Nos. 08/853,826 and 09/074,720, the disclosures of which are incorporated herein by reference in their entireties.

Any vaccine provided herein may be prepared using well known methods that result in a combination of antigen, immune response enhancer and a suitable carrier or excipient. The compositions described herein may be administered as part of a sustained release formulation (*i.e.*, a formulation such as a capsule, sponge or gel (composed of polysaccharides, for example) that effects a slow release of compound following administration). Such formulations may generally be prepared using well known technology (*see, e.g.*, Coombes et al., *Vaccine* 14:1429-1438, 1996) and administered by, for example, oral, rectal or subcutaneous implantation, or by

implantation at the desired target site. Sustained-release formulations may contain a polypeptide, polynucleotide or antibody dispersed in a carrier matrix and/or contained within a reservoir surrounded by a rate controlling membrane.

Carriers for use within such formulations are biocompatible, and may 5 also be biodegradable; preferably the formulation provides a relatively constant level of active component release. Such carriers include microparticles of poly(lactide-co-glycolide), polyacrylate, latex, starch, cellulose, dextran and the like. Other delayed-release carriers include supramolecular biovectors, which comprise a non-liquid hydrophilic core (e.g., a cross-linked polysaccharide or oligosaccharide) and, optionally, 10 an external layer comprising an amphiphilic compound, such as a phospholipid (see e.g., U.S. Patent No. 5,151,254 and PCT applications WO 94/20078, WO/94/23701 and WO 96/06638). The amount of active compound contained within a sustained release formulation depends upon the site of implantation, the rate and expected duration of release and the nature of the condition to be treated or prevented.

15 Any of a variety of delivery vehicles may be employed within pharmaceutical compositions and vaccines to facilitate production of an antigen-specific immune response that targets tumor cells. Delivery vehicles include antigen presenting cells (APCs), such as dendritic cells, macrophages, B cells, monocytes and other cells that may be engineered to be efficient APCs. Such cells may, but need not, be 20 genetically modified to increase the capacity for presenting the antigen, to improve activation and/or maintenance of the T cell response, to have anti-tumor effects *per se* and/or to be immunologically compatible with the receiver (i.e., matched HLA haplotype). APCs may generally be isolated from any of a variety of biological fluids and organs, including tumor and peritumoral tissues, and may be autologous, 25 allogeneic, syngeneic or xenogeneic cells.

Certain preferred embodiments of the present invention use dendritic cells or progenitors thereof as antigen-presenting cells. Dendritic cells are highly potent APCs (Banchereau and Steinman, *Nature* 392:245-251, 1998) and have been shown to be effective as a physiological adjuvant for eliciting prophylactic or therapeutic 30 antitumor immunity (see Timmerman and Levy, *Ann. Rev. Med.* 50:507-529, 1999). In general, dendritic cells may be identified based on their typical shape (stellate *in situ*,

with marked cytoplasmic processes (dendrites) visible *in vitro*), their ability to take up, process and present antigens with high efficiency and their ability to activate naïve T cell responses. Dendritic cells may, of course, be engineered to express specific cell-surface receptors or ligands that are not commonly found on dendritic cells *in vivo* or *ex vivo*, and such modified dendritic cells are contemplated by the present invention. As an alternative to dendritic cells, secreted vesicles antigen-loaded dendritic cells (called exosomes) may be used within a vaccine (see Zitvogel et al., *Nature Med.* 4:594-600, 1998).

Dendritic cells and progenitors may be obtained from peripheral blood, bone marrow, tumor-infiltrating cells, peritumoral tissues-infiltrating cells, lymph nodes, spleen, skin, umbilical cord blood or any other suitable tissue or fluid. For example, dendritic cells may be differentiated *ex vivo* by adding a combination of cytokines such as GM-CSF, IL-4, IL-13 and/or TNF α to cultures of monocytes harvested from peripheral blood. Alternatively, CD34 positive cells harvested from peripheral blood, umbilical cord blood or bone marrow may be differentiated into dendritic cells by adding to the culture medium combinations of GM-CSF, IL-3, TNF α , CD40 ligand, LPS, flt3 ligand and/or other compound(s) that induce differentiation, maturation and proliferation of dendritic cells.

Dendritic cells are conveniently categorized as "immature" and "mature" cells, which allows a simple way to discriminate between two well characterized phenotypes. However, this nomenclature should not be construed to exclude all possible intermediate stages of differentiation. Immature dendritic cells are characterized as APC with a high capacity for antigen uptake and processing, which correlates with the high expression of Fc γ receptor and mannose receptor. The mature phenotype is typically characterized by a lower expression of these markers, but a high expression of cell surface molecules responsible for T cell activation such as class I and class II MHC, adhesion molecules (e.g., CD54 and CD11) and costimulatory molecules (e.g., CD40, CD80, CD86 and 4-1BB).

APCs may generally be transfected with a polynucleotide encoding a breast tumor protein (or portion or other variant thereof) such that the breast tumor polypeptide, or an immunogenic portion thereof, is expressed on the cell surface. Such

transfection may take place *ex vivo*, and a composition or vaccine comprising such transfected cells may then be used for therapeutic purposes, as described herein. Alternatively, a gene delivery vehicle that targets a dendritic or other antigen presenting cell may be administered to a patient, resulting in transfection that occurs *in vivo*. *In vivo* and *ex vivo* transfection of dendritic cells, for example, may generally be performed using any methods known in the art, such as those described in WO 5 97/24447, or the gene gun approach described by Mahvi et al., *Immunology and cell Biology* 75:456-460, 1997. Antigen loading of dendritic cells may be achieved by incubating dendritic cells or progenitor cells with the breast tumor polypeptide, DNA 10 (naked or within a plasmid vector) or RNA; or with antigen-expressing recombinant bacterium or viruses (e.g., vaccinia, fowlpox, adenovirus or lentivirus vectors). Prior to loading, the polypeptide may be covalently conjugated to an immunological partner that provides T cell help (e.g., a carrier molecule). Alternatively, a dendritic cell may be 15 pulsed with a non-conjugated immunological partner, separately or in the presence of the polypeptide.

Vaccines and pharmaceutical compositions may be presented in unit-dose or multi-dose containers, such as sealed ampoules or vials. Such containers are 20 preferably hermetically sealed to preserve sterility of the formulation until use. In general, formulations may be stored as suspensions, solutions or emulsions in oily or aqueous vehicles. Alternatively, a vaccine or pharmaceutical composition may be stored in a freeze-dried condition requiring only the addition of a sterile liquid carrier immediately prior to use.

CANCER THERAPY

In further aspects of the present invention, the compositions described 25 herein may be used for immunotherapy of cancer, such as breast cancer. Within such methods, pharmaceutical compositions and vaccines are typically administered to a patient. As used herein, a "patient" refers to any warm-blooded animal, preferably a human. A patient may or may not be afflicted with cancer. Accordingly, the above pharmaceutical compositions and vaccines may be used to prevent the development of a 30 cancer or to treat a patient afflicted with a cancer. A cancer may be diagnosed using

criteria generally accepted in the art, including the presence of a malignant tumor. Pharmaceutical compositions and vaccines may be administered either prior to or following surgical removal of primary tumors and/or treatment such as administration of radiotherapy or conventional chemotherapeutic drugs. Administration may be by any 5 suitable method, including administration by intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal, intradermal, anal, vaginal, topical and oral routes.

Within certain embodiments, immunotherapy may be active immunotherapy, in which treatment relies on the *in vivo* stimulation of the endogenous 10 host immune system to react against tumors with the administration of immune response-modifying agents (such as polypeptides and polynucleotides as provided herein).

Within other embodiments, immunotherapy may be passive immunotherapy, in which treatment involves the delivery of agents with established 15 tumor-immune reactivity (such as effector cells or antibodies) that can directly or indirectly mediate antitumor effects and does not necessarily depend on an intact host immune system. Examples of effector cells include T cells as discussed above, T lymphocytes (such as CD8⁺ cytotoxic T lymphocytes and CD4⁺ T-helper tumor-infiltrating lymphocytes), killer cells (such as Natural Killer cells and lymphokine- 20 activated killer cells), B cells and antigen-presenting cells (such as dendritic cells and macrophages) expressing a polypeptide provided herein. T cell receptors and antibody receptors specific for the polypeptides recited herein may be cloned, expressed and transferred into other vectors or effector cells for adoptive immunotherapy. The polypeptides provided herein may also be used to generate antibodies or anti-idiotypic 25 antibodies (as described above and in U.S. Patent No. 4,918,164) for passive immunotherapy.

Effector cells may generally be obtained in sufficient quantities for adoptive immunotherapy by growth *in vitro*, as described herein. Culture conditions for expanding single antigen-specific effector cells to several billion in number with 30 retention of antigen recognition *in vivo* are well known in the art. Such *in vitro* culture conditions typically use intermittent stimulation with antigen, often in the presence of

cytokines (such as IL-2) and non-dividing feeder cells. As noted above, immunoreactive polypeptides as provided herein may be used to rapidly expand antigen-specific T cell cultures in order to generate a sufficient number of cells for immunotherapy. In particular, antigen-presenting cells, such as dendritic, macrophage, 5 monocyte, fibroblast and/or B cells, may be pulsed with immunoreactive polypeptides or transfected with one or more polynucleotides using standard techniques well known in the art. For example, antigen-presenting cells can be transfected with a polynucleotide having a promoter appropriate for increasing expression in a recombinant virus or other expression system. Cultured effector cells for use in therapy 10 must be able to grow and distribute widely, and to survive long term *in vivo*. Studies have shown that cultured effector cells can be induced to grow *in vivo* and to survive long term in substantial numbers by repeated stimulation with antigen supplemented with IL-2 (see, for example, Cheever et al., *Immunological Reviews* 157:177, 1997).

15 Alternatively, a vector expressing a polypeptide recited herein may be introduced into antigen presenting cells taken from a patient and clonally propagated *ex vivo* for transplant back into the same patient. Transfected cells may be reintroduced into the patient using any means known in the art, preferably in sterile form by intravenous, intracavitory, intraperitoneal or intratumor administration.

Routes and frequency of administration of the therapeutic compositions 20 described herein, as well as dosage, will vary from individual to individual, and may be readily established using standard techniques. In general, the pharmaceutical compositions and vaccines may be administered by injection (e.g., intracutaneous, intramuscular, intravenous or subcutaneous), intranasally (e.g., by aspiration) or orally. Preferably, between 1 and 10 doses may be administered over a 52 week period. 25 Preferably, 6 doses are administered, at intervals of 1 month, and booster vaccinations may be given periodically thereafter. Alternate protocols may be appropriate for individual patients. A suitable dose is an amount of a compound that, when administered as described above, is capable of promoting an anti-tumor immune response, and is at least 10-50% above the basal (i.e., untreated) level. Such response 30 can be monitored by measuring the anti-tumor antibodies in a patient or by vaccine-dependent generation of cytolytic effector cells capable of killing the patient's tumor

cells *in vitro*. Such vaccines should also be capable of causing an immune response that leads to an improved clinical outcome (e.g., more frequent remissions, complete or partial or longer disease-free survival) in vaccinated patients as compared to non-vaccinated patients. In general, for pharmaceutical compositions and vaccines 5 comprising one or more polypeptides, the amount of each polypeptide present in a dose ranges from about 25 µg to 5 mg per kg of host. Suitable dose sizes will vary with the size of the patient, but will typically range from about 0.1 mL to about 5 mL.

In general, an appropriate dosage and treatment regimen provides the active compound(s) in an amount sufficient to provide therapeutic and/or prophylactic 10 benefit. Such a response can be monitored by establishing an improved clinical outcome (e.g., more frequent remissions, complete or partial, or longer disease-free survival) in treated patients as compared to non-treated patients. Increases in preexisting immune responses to a breast tumor protein generally correlate with an improved clinical outcome. Such immune responses may generally be evaluated using 15 standard proliferation, cytotoxicity or cytokine assays, which may be performed using samples obtained from a patient before and after treatment.

CANCER DETECTION AND DIAGNOSIS

In general, a cancer may be detected in a patient based on the presence of one or more breast tumor proteins and/or polynucleotides encoding such proteins in a 20 biological sample (for example, blood, sera, sputum urine and/or tumor biopsies) obtained from the patient. In other words, such proteins may be used as markers to indicate the presence or absence of a cancer such as breast cancer. In addition, such proteins may be useful for the detection of other cancers. The binding agents provided herein generally permit detection of the level of antigen that binds to the agent in the 25 biological sample. Polynucleotide primers and probes may be used to detect the level of mRNA encoding a tumor protein, which is also indicative of the presence or absence of a cancer. In general, a breast tumor sequence should be present at a level that is at least three fold higher in tumor tissue than in normal tissue

There are a variety of assay formats known to those of ordinary skill in 30 the art for using a binding agent to detect polypeptide markers in a sample. *See, e.g.*,

Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, the presence or absence of a cancer in a patient may be determined by (a) contacting a biological sample obtained from a patient with a binding agent; (b) detecting in the sample a level of polypeptide that binds to the binding agent; and (c) 5 comparing the level of polypeptide with a predetermined cut-off value.

In a preferred embodiment, the assay involves the use of binding agent immobilized on a solid support to bind to and remove the polypeptide from the remainder of the sample. The bound polypeptide may then be detected using a detection reagent that contains a reporter group and specifically binds to the binding 10 agent/polypeptide complex. Such detection reagents may comprise, for example, a binding agent that specifically binds to the polypeptide or an antibody or other agent that specifically binds to the binding agent, such as an anti-immunoglobulin, protein G, protein A or a lectin. Alternatively, a competitive assay may be utilized, in which a polypeptide is labeled with a reporter group and allowed to bind to the immobilized 15 binding agent after incubation of the binding agent with the sample. The extent to which components of the sample inhibit the binding of the labeled polypeptide to the binding agent is indicative of the reactivity of the sample with the immobilized binding agent. Suitable polypeptides for use within such assays include full length breast tumor proteins and portions thereof to which the binding agent binds, as described above.

20 The solid support may be any material known to those of ordinary skill in the art to which the tumor protein may be attached. For example, the solid support may be a test well in a microtiter plate or a nitrocellulose or other suitable membrane. Alternatively, the support may be a bead or disc, such as glass, fiberglass, latex or a plastic material such as polystyrene or polyvinylchloride. The support may also be a 25 magnetic particle or a fiber optic sensor, such as those disclosed, for example, in U.S. Patent No. 5,359,681. The binding agent may be immobilized on the solid support using a variety of techniques known to those of skill in the art, which are amply described in the patent and scientific literature. In the context of the present invention, the term "immobilization" refers to both noncovalent association, such as adsorption, 30 and covalent attachment (which may be a direct linkage between the agent and functional groups on the support or may be a linkage by way of a cross-linking agent).

Immobilization by adsorption to a well in a microtiter plate or to a membrane is preferred. In such cases, adsorption may be achieved by contacting the binding agent, in a suitable buffer, with the solid support for a suitable amount of time. The contact time varies with temperature, but is typically between about 1 hour and about 1 day. In 5 general, contacting a well of a plastic microtiter plate (such as polystyrene or polyvinylchloride) with an amount of binding agent ranging from about 10 ng to about 10 μ g, and preferably about 100 ng to about 1 μ g, is sufficient to immobilize an adequate amount of binding agent.

Covalent attachment of binding agent to a solid support may generally be 10 achieved by first reacting the support with a bifunctional reagent that will react with both the support and a functional group, such as a hydroxyl or amino group, on the binding agent. For example, the binding agent may be covalently attached to supports having an appropriate polymer coating using benzoquinone or by condensation of an aldehyde group on the support with an amine and an active hydrogen on the binding 15 partner (see, e.g., Pierce Immunotechnology Catalog and Handbook, 1991, at A12-A13).

In certain embodiments, the assay is a two-antibody sandwich assay. This assay may be performed by first contacting an antibody that has been immobilized 20 on a solid support, commonly the well of a microtiter plate, with the sample, such that polypeptides within the sample are allowed to bind to the immobilized antibody. Unbound sample is then removed from the immobilized polypeptide-antibody complexes and a detection reagent (preferably a second antibody capable of binding to a different site on the polypeptide) containing a reporter group is added. The amount of detection reagent that remains bound to the solid support is then determined using a 25 method appropriate for the specific reporter group.

More specifically, once the antibody is immobilized on the support as described above, the remaining protein binding sites on the support are typically blocked. Any suitable blocking agent known to those of ordinary skill in the art, such as bovine serum albumin or Tween 20TM (Sigma Chemical Co., St. Louis, MO). The 30 immobilized antibody is then incubated with the sample, and polypeptide is allowed to bind to the antibody. The sample may be diluted with a suitable diluent, such as

phosphate-buffered saline (PBS) prior to incubation. In general, an appropriate contact time (*i.e.*, incubation time) is a period of time that is sufficient to detect the presence of polypeptide within a sample obtained from an individual with breast cancer. Preferably, the contact time is sufficient to achieve a level of binding that is at least about 95% of 5 that achieved at equilibrium between bound and unbound polypeptide. Those of ordinary skill in the art will recognize that the time necessary to achieve equilibrium may be readily determined by assaying the level of binding that occurs over a period of time. At room temperature, an incubation time of about 30 minutes is generally sufficient.

10 Unbound sample may then be removed by washing the solid support with an appropriate buffer, such as PBS containing 0.1% Tween 20TM. The second antibody, which contains a reporter group, may then be added to the solid support. Preferred reporter groups include those groups recited above.

15 The detection reagent is then incubated with the immobilized antibody-polypeptide complex for an amount of time sufficient to detect the bound polypeptide. An appropriate amount of time may generally be determined by assaying the level of binding that occurs over a period of time. Unbound detection reagent is then removed and bound detection reagent is detected using the reporter group. The method employed for detecting the reporter group depends upon the nature of the reporter 20 group. For radioactive groups, scintillation counting or autoradiographic methods are generally appropriate. Spectroscopic methods may be used to detect dyes, luminescent groups and fluorescent groups. Biotin may be detected using avidin, coupled to a different reporter group (commonly a radioactive or fluorescent group or an enzyme). Enzyme reporter groups may generally be detected by the addition of substrate 25 (generally for a specific period of time), followed by spectroscopic or other analysis of the reaction products.

30 To determine the presence or absence of a cancer, such as breast cancer, the signal detected from the reporter group that remains bound to the solid support is generally compared to a signal that corresponds to a predetermined cut-off value. In one preferred embodiment, the cut-off value for the detection of a cancer is the average mean signal obtained when the immobilized antibody is incubated with samples from

patients without the cancer. In general, a sample generating a signal that is three standard deviations above the predetermined cut-off value is considered positive for the cancer. In an alternate preferred embodiment, the cut-off value is determined using a Receiver Operator Curve, according to the method of Sackett et al., *Clinical Epidemiology: A Basic Science for Clinical Medicine*, Little Brown and Co., 1985, p. 106-7. Briefly, in this embodiment, the cut-off value may be determined from a plot of pairs of true positive rates (*i.e.*, sensitivity) and false positive rates (100%-specificity) that correspond to each possible cut-off value for the diagnostic test result. The cut-off value on the plot that is the closest to the upper left-hand corner (*i.e.*, the value that encloses the largest area) is the most accurate cut-off value, and a sample generating a signal that is higher than the cut-off value determined by this method may be considered positive. Alternatively, the cut-off value may be shifted to the left along the plot, to minimize the false positive rate, or to the right, to minimize the false negative rate. In general, a sample generating a signal that is higher than the cut-off value determined by this method is considered positive for a cancer.

In a related embodiment, the assay is performed in a flow-through or strip test format, wherein the binding agent is immobilized on a membrane, such as nitrocellulose. In the flow-through test, polypeptides within the sample bind to the immobilized binding agent as the sample passes through the membrane. A second, labeled binding agent then binds to the binding agent-polypeptide complex as a solution containing the second binding agent flows through the membrane. The detection of bound second binding agent may then be performed as described above. In the strip test format, one end of the membrane to which binding agent is bound is immersed in a solution containing the sample. The sample migrates along the membrane through a region containing second binding agent and to the area of immobilized binding agent. Concentration of second binding agent at the area of immobilized antibody indicates the presence of a cancer. Typically, the concentration of second binding agent at that site generates a pattern, such as a line, that can be read visually. The absence of such a pattern indicates a negative result. In general, the amount of binding agent immobilized on the membrane is selected to generate a visually discernible pattern when the biological sample contains a level of polypeptide that would be sufficient to generate a

positive signal in the two-antibody sandwich assay, in the format discussed above. Preferred binding agents for use in such assays are antibodies and antigen-binding fragments thereof. Preferably, the amount of antibody immobilized on the membrane ranges from about 25 ng to about 1 μ g, and more preferably from about 50 ng to about 5 500 ng. Such tests can typically be performed with a very small amount of biological sample.

Of course, numerous other assay protocols exist that are suitable for use with the tumor proteins or binding agents of the present invention. The above descriptions are intended to be exemplary only. For example, it will be apparent to 10 those of ordinary skill in the art that the above protocols may be readily modified to use breast tumor polypeptides to detect antibodies that bind to such polypeptides in a biological sample. The detection of such breast tumor protein specific antibodies may correlate with the presence of a cancer.

A cancer may also, or alternatively, be detected based on the presence of 15 T cells that specifically react with a breast tumor protein in a biological sample. Within certain methods, a biological sample comprising CD4 $^{+}$ and/or CD8 $^{+}$ T cells isolated from a patient is incubated with a breast tumor polypeptide, a polynucleotide encoding such a polypeptide and/or an APC that expresses at least an immunogenic portion of such a polypeptide, and the presence or absence of specific activation of the T cells is 20 detected. Suitable biological samples include, but are not limited to, isolated T cells. For example, T cells may be isolated from a patient by routine techniques (such as by Ficoll/Hypaque density gradient centrifugation of peripheral blood lymphocytes). T cells may be incubated *in vitro* for 2-9 days (typically 4 days) at 37°C with polypeptide (e.g., 5 - 25 μ g/ml). It may be desirable to incubate another aliquot of a T cell sample in 25 the absence of breast tumor polypeptide to serve as a control. For CD4 $^{+}$ T cells, activation is preferably detected by evaluating proliferation of the T cells. For CD8 $^{+}$ T cells, activation is preferably detected by evaluating cytolytic activity. A level of proliferation that is at least two fold greater and/or a level of cytolytic activity that is at least 20% greater than in disease-free patients indicates the presence of a cancer in the 30 patient.

As noted above, a cancer may also, or alternatively, be detected based on the level of mRNA encoding a breast tumor protein in a biological sample. For example, at least two oligonucleotide primers may be employed in a polymerase chain reaction (PCR) based assay to amplify a portion of a breast tumor cDNA derived from a 5 biological sample, wherein at least one of the oligonucleotide primers is specific for (*i.e.*, hybridizes to) a polynucleotide encoding the breast tumor protein. The amplified cDNA is then separated and detected using techniques well known in the art, such as gel electrophoresis. Similarly, oligonucleotide probes that specifically hybridize to a polynucleotide encoding a breast tumor protein may be used in a hybridization assay to 10 detect the presence of polynucleotide encoding the tumor protein in a biological sample.

To permit hybridization under assay conditions, oligonucleotide primers and probes should comprise an oligonucleotide sequence that has at least about 60%, preferably at least about 75% and more preferably at least about 90%, identity to a portion of a polynucleotide encoding a breast tumor protein that is at least 10 15 nucleotides, and preferably at least 20 nucleotides, in length. Preferably, oligonucleotide primers and/or probes hybridize to a polynucleotide encoding a polypeptide described herein under moderately stringent conditions, as defined above. Oligonucleotide primers and/or probes which may be usefully employed in the diagnostic methods described herein preferably are at least 10-40 nucleotides in length. 20 In a preferred embodiment, the oligonucleotide primers comprise at least 10 contiguous nucleotides, more preferably at least 15 contiguous nucleotides, of a DNA molecule having a sequence recited in SEQ ID NO: 1-38, 42-204, 205, 207 and 210-290. Techniques for both PCR based assays and hybridization assays are well known in the art (*see, for example, Mullis et al., Cold Spring Harbor Symp. Quant. Biol., 51:263, 25 1987; Erlich ed., PCR Technology, Stockton Press, NY, 1989.*).

One preferred assay employs RT-PCR, in which PCR is applied in conjunction with reverse transcription. Typically, RNA is extracted from a biological sample, such as biopsy tissue, and is reverse transcribed to produce cDNA molecules. PCR amplification using at least one specific primer generates a cDNA molecule, which 30 may be separated and visualized using, for example, gel electrophoresis. Amplification may be performed on biological samples taken from a test patient and from an

individual who is not afflicted with a cancer. The amplification reaction may be performed on several dilutions of cDNA spanning two orders of magnitude. A two-fold or greater increase in expression in several dilutions of the test patient sample as compared to the same dilutions of the non-cancerous sample is typically considered 5 positive.

In another embodiment, the compositions described herein may be used as markers for the progression of cancer. In this embodiment, assays as described above for the diagnosis of a cancer may be performed over time, and the change in the level of reactive polypeptide(s) or polynucleotide(s) evaluated. For example, the assays 10 may be performed every 24-72 hours for a period of 6 months to 1 year, and thereafter performed as needed. In general, a cancer is progressing in those patients in whom the level of polypeptide or polynucleotide detected increases over time. In contrast, the cancer is not progressing when the level of reactive polypeptide or polynucleotide either remains constant or decreases with time.

15 Certain *in vivo* diagnostic assays may be performed directly on a tumor. One such assay involves contacting tumor cells with a binding agent. The bound binding agent may then be detected directly or indirectly via a reporter group. Such binding agents may also be used in histological applications. Alternatively, polynucleotide probes may be used within such applications.

20 As noted above, to improve sensitivity, multiple breast tumor protein markers may be assayed within a given sample. It will be apparent that binding agents specific for different proteins provided herein may be combined within a single assay. Further, multiple primers or probes may be used concurrently. The selection of tumor protein markers may be based on routine experiments to determine combinations that 25 results in optimal sensitivity. In addition, or alternatively, assays for tumor proteins provided herein may be combined with assays for other known tumor antigens.

DIAGNOSTIC KITS

The present invention further provides kits for use within any of the above diagnostic methods. Such kits typically comprise two or more components 30 necessary for performing a diagnostic assay. Components may be compounds,

reagents, containers and/or equipment. For example, one container within a kit may contain a monoclonal antibody or fragment thereof that specifically binds to a breast tumor protein. Such antibodies or fragments may be provided attached to a support material, as described above. One or more additional containers may enclose elements, 5 such as reagents or buffers, to be used in the assay. Such kits may also, or alternatively, contain a detection reagent as described above that contains a reporter group suitable for direct or indirect detection of antibody binding.

Alternatively, a kit may be designed to detect the level of mRNA encoding a breast tumor protein in a biological sample. Such kits generally comprise at 10 least one oligonucleotide probe or primer, as described above, that hybridizes to a polynucleotide encoding a breast tumor protein. Such an oligonucleotide may be used, for example, within a PCR or hybridization assay. Additional components that may be present within such kits include a second oligonucleotide and/or a diagnostic reagent or container to facilitate the detection of a polynucleotide encoding a breast tumor protein.

15 The following Examples are offered by way of illustration and not by way of limitation.

EXAMPLE 1

IDENTIFICATION OF BREAST TUMOR PROTEIN cDNAS USING
SUBTRACTION METHODOLOGY

This Example illustrates the identification of cDNA molecules encoding
5 breast tumor proteins.

A human metastatic breast tumor cDNA expression library was constructed from metastatic breast tumor poly A⁺ RNA using a Superscript Plasmid System for cDNA Synthesis and Plasmid Cloning kit (BRL Life Technologies, Gaithersburg, MD 20897) following the manufacturer's protocol. Specifically, breast
10 tumor tissues were homogenized with polytron (Kinematica, Switzerland) and total RNA was extracted using Trizol reagent (BRL Life Technologies) as directed by the manufacturer. The poly A⁺ RNA was then purified using a Qiagen oligotex spin column mRNA purification kit (Qiagen, Santa Clarita, CA 91355) according to the manufacturer's protocol. First-strand cDNA was synthesized using the NotI/Oligo-
15 dT18 primer. Double-stranded cDNA was synthesized, ligated with EcoRI/BstX I adaptors (Invitrogen, Carlsbad, CA) and digested with NotI. Following size fractionation with Chroma Spin-1000 columns (Clontech, Palo Alto, CA 94303), the cDNA was ligated into the EcoRI/NotI site of pCDNA3.1 (Invitrogen, Carlsbad, CA) and transformed into ElectroMax *E. coli* DH10B cells (BRL Life Technologies) by
20 electroporation.

Using the same procedure, a normal human breast cDNA expression library was prepared from a pool of four normal breast tissue specimens. The cDNA libraries were characterized by determining the number of independent colonies, the percentage of clones that carried insert, the average insert size and by sequence analysis.
25 Sequencing analysis showed both libraries to contain good complex cDNA clones that were synthesized from mRNA, with minimal rRNA and mitochondrial DNA contamination sequencing.

A cDNA subtracted library (referred to as BS3) was prepared using the above metastatic breast tumor and normal breast cDNA libraries, as described by Hara
30 *et al.* (*Blood*, 84:189-199, 1994) with some modifications. Specifically, a breast tumor-

specific subtracted cDNA library was generated as follows. Normal breast cDNA library (70 µg) was digested with EcoRI, NotI, and SfI, followed by a filling-in reaction with DNA polymerase Klenow fragment. After phenol-chloroform extraction and ethanol precipitation, the DNA was dissolved in 100 µl of H₂O, heat-denatured and 5 mixed with 100 µl (100 µg) of Photoprobe biotin (Vector Laboratories, Burlingame, CA), the resulting mixture was irradiated with a 270 W sunlamp on ice for 20 minutes. Additional Photoprobe biotin (50 µl) was added and the biotinylation reaction was repeated. After extraction with butanol five times, the DNA was ethanol-precipitated and dissolved in 23 µl H₂O to form the driver DNA.

10 To form the tracer DNA, 10 µg breast tumor cDNA library was digested with BamHI and XhoI, phenol chloroform extracted and passed through Chroma spin-400 columns (Clontech). Following ethanol precipitation, the tracer DNA was dissolved in 5 µl H₂O. Tracer DNA was mixed with 15 µl driver DNA and 20 µl of 2 x hybridization buffer (1.5 M NaCl/10 mM EDTA/50 mM HEPES pH 7.5/0.2% sodium 15 dodecyl sulfate), overlaid with mineral oil, and heat-denatured completely. The sample was immediately transferred into a 68 °C water bath and incubated for 20 hours (long hybridization [LH]). The reaction mixture was then subjected to a streptavidin treatment followed by phenol/chloroform extraction. This process was repeated three more times. Subtracted DNA was precipitated, dissolved in 12 µl H₂O, mixed with 8 µl 20 driver DNA and 20 µl of 2 x hybridization buffer, and subjected to a hybridization at 68 °C for 2 hours (short hybridization [SH]). After removal of biotinylated double-stranded DNA, subtracted cDNA was ligated into BamHI/XhoI site of chloramphenicol 25 resistant pBCSK⁺ (Stratagene, La Jolla, CA 92037) and transformed into ElectroMax *E. coli* DH10B cells by electroporation to generate a breast tumor specific subtracted cDNA library.

To analyze the subtracted cDNA library, plasmid DNA was prepared from independent clones, randomly picked from the subtracted breast tumor specific library and characterized by DNA sequencing with a Perkin Elmer/Applied Biosystems Division Automated Sequencer Model 373A (Foster City, CA).

30 A second cDNA subtraction library containing cDNA from breast tumor subtracted with normal breast cDNA, and known as BT, was constructed as follows.

Total RNA was extracted from primary breast tumor tissues using Trizol reagent (Gibco BRL Life Technologies, Gaithersburg, MD) as described by the manufacturer. The polyA+ RNA was purified using an oligo(dT) cellulose column according to standard protocols. First strand cDNA was synthesized using the primer supplied in a Clontech 5 PCR-Select cDNA Subtraction Kit (Clontech, Palo Alto, CA). The driver DNA consisted of cDNAs from two normal breast tissues with the tester cDNA being from three primary breast tumors. Double-stranded cDNA was synthesized for both tester and driver, and digested with a combination of endonucleases (MluI, MscI, PvuII, SalI and StuI) which recognize six base pairs DNA. This modification increased the average 10 cDNA size dramatically compared with cDNAs generated according to the protocol of Clontech. The digested tester cDNAs were ligated to two different adaptors and the subtraction was performed according to Clontech's protocol. The subtracted cDNAs were subjected to two rounds of PCR amplification, following the manufacturer's protocol. The resulting PCR products were subcloned into the TA cloning vector, 15 pCRII (Invitrogen, San Diego, CA) and transformed into ElectroMax *E. coli* DH10B cells (Gibco BRL Life, Technologies) by electroporation. DNA was isolated from independent clones and sequenced using a Perkin Elmer/Applied Biosystems Division (Foster City, CA) Automated Sequencer Model 373A.

Two additional subtracted cDNA libraries were prepared from cDNA 20 from breast tumors subtracted with a pool of cDNA from six normal tissues (liver, brain, stomach, small intestine, kidney and heart; referred to as 2BT and BC6) using the PCR-subtraction protocol of Clontech, described above. A fourth subtracted library (referred to as Bt-Met) was prepared using the protocol of Clontech from cDNA from metastatic breast tumors subtracted with cDNA from five normal tissues (brain, lung, 25 PBMC, pancreas and normal breast).

cDNA clones isolated in the breast subtractions BS3, BT, 2BT, BC6 and BT-Met, described above, were colony PCR amplified and their mRNA expression levels in breast tumor, normal breast and various other normal tissues were determined using microarray technology. Briefly, the PCR amplification products were dotted onto 30 slides in an array format, with each product occupying a unique location in the array. mRNA was extracted from the tissue sample to be tested, reverse transcribed, and

fluorescent-labeled cDNA probes were generated. The microarrays were probed with the labeled cDNA probes, the slides scanned and fluorescence intensity was measured. This intensity correlates with the hybridization intensity.

The determined cDNA sequences of 131 clones determined to be over-expressed in breast tumor tissue compared to other tissues tested by a visual analysis of the microarray data are provided in SEQ ID NO: 1-35 and 42-137. Comparison of these cDNA sequences with known sequences in the gene bank using the EMBL and GenBank databases revealed no significant homologies to the sequences provided in SEQ ID NO: 7, 10, 21, 26, 30, 63, 81 and 104. The sequences of SEQ ID NO: 2-5, 8, 9, 10, 13, 15, 16, 22, 25, 27, 28, 33, 35, 72, 73, 103, 107, 109, 118, 128, 129 134 and 136 showed some homology to previously isolated expressed sequences tags (ESTs), while the sequences of SEQ ID NO: 1, 6, 11, 12, 14, 17-20, 23, 24, 29, 31, 32, 34, 42-62, 64-71, 74-80, 82-102, 105, 106, 108, 110-117, 119-127, 130-133, 135 and 137 showed some homology to previously identified genes.

15 The determined cDNA sequences of an additional 45 clones isolated from the BT-Met library as described above and found to be over-expressed in breast tumors and metastatic breast tumors compared to other tissues tested, are provided in SEQ ID NO: 138-182. Comparison of the sequences of SEQ ID NO: 159-161, 164 and 181 revealed no significant homologies to previously identified sequences. The 20 sequences of SEQ ID NO: 138-158, 162, 163, 165-180 and 182 showed some homology to previously identified genes.

In further studies, suppression subtractive hybridization (Clontech) was preformed using a pool of cDNA from 3 unique human breast tumors as the tester and a pool of cDNA from 6 other normal human tissues (liver, brain, stomach, small intestine, 25 heart and kidney) as the driver. The isolated cDNA fragments were subcloned and characterized by DNA sequencing. The determined cDNA sequences of 22 isolated clones are provided in SEQ ID NO: 183-204. Comparison of these sequences with those in the public databases revealed no significant homologies to previously identified sequences.

30 The determined cDNA sequences of 71 additional breast-specific genes isolated during characterization of breast tumor cDNA libraries are provided in SEQ ID

NO: 210-290. Comparison of these sequences with those in the GenBank and Geneseq databases revealed no significant homologies.

EXAMPLE 2

5 IDENTIFICATION OF BREAST TUMOR PROTEIN cDNAs BY RT-PCR

GABA_A receptor clones were isolated from human breast cancer cDNA libraries by first preparing cDNA libraries from breast tumor samples from different patients as described above. PCR primers were designed based on the GABA_A receptor subunit sequences described by Hedblom and Kirkness (*Jnl. Biol. Chem.* 272:15346-10 15350, 1997) and used to amplify sequences from the breast tumor cDNA libraries by RT-PCR. The determined cDNA sequences of three GABA_A receptor clones are provided in SEQ ID NO: 36-38, with the corresponding amino acid sequences being provided in SEQ ID NO: 39-41.

15 The clone with the longest open reading frame (ORF; SEQ ID NO: 36) showed homology to the GABA_A receptor of Hedblom and Kirkness, with four potential transmembrane regions at the C-terminal part of the protein, while the clones of SEQ ID NO: 37 and 38 retained either no transmembrane region or only the first transmembrane region. Some patients were found to have only the clones with the shorter ORFs while others had both the clones with longer and shorter ORFs.

20

EXAMPLE 3

EXPRESSION OF OVARIAN TUMOR-DERIVED ANTIGENS
IN BREAST

25 Isolation of the antigens O772P and O8E from ovarian tumor tissue is described in US Patent Application No. 09/338,933, filed June 23, 1999. The determined cDNA sequence for O772P is provided in SEQ ID NO: 205, with the corresponding amino acid sequence being provided in SEQ ID NO: 206. The full-length cDNA sequence for O8E is provided in SEQ ID NO: 207. Two protein sequences may be translated from the full length O8E. Form "A" (SEQ ID NO: 208)

begins with a putative start methionine. A second form "B" (SEQ ID NO: 209) includes 27 additional upstream residues to the 5' end of the nucleotide sequence.

The expression levels of O772P and O8E in a variety of tumor and normal tissues, including metastatic breast tumors, were analyzed by real time PCR.

5 Both genes were found to have increased mRNA expression in 30-50% of breast tumors. For O772P, elevated expression was also observed in normal trachea, ureter, uterus and ovary. For O8E, elevated expression was also observed in normal trachea, kidney and ovary. Additional analysis employing a panel of tumor cell lines demonstrated increased expression of O8E in the breast tumor cell lines SKBR3, MDA-
10 MB-415 and BT474, and increased expression of O772P in SKBR3. Collectively, the data indicate that O772P and O8E may be useful in the diagnosis and therapy of breast cancer.

EXAMPLE 4

15 SYNTHESIS OF POLYPEPTIDES

Polypeptides may be synthesized on a Perkin Elmer/Applied Biosystems Division 430A peptide synthesizer using FMOC chemistry with HPTU (O-Benzotriazole-N,N,N',N'-tetramethyluronium hexafluorophosphate) activation. A Gly-Cys-Gly sequence may be attached to the amino terminus of the peptide to provide a
20 method of conjugation, binding to an immobilized surface, or labeling of the peptide. Cleavage of the peptides from the solid support may be carried out using the following cleavage mixture: trifluoroacetic acid:ethanedithiol:thioanisole:water:phenol (40:1:2:2:3). After cleaving for 2 hours, the peptides may be precipitated in cold methyl-t-butyl-ether. The peptide pellets may then be dissolved in water containing
25 0.1% trifluoroacetic acid (TFA) and lyophilized prior to purification by C18 reverse phase HPLC. A gradient of 0%-60% acetonitrile (containing 0.1% TFA) in water (containing 0.1% TFA) may be used to elute the peptides. Following lyophilization of the pure fractions, the peptides may be characterized using electrospray or other types of mass spectrometry and by amino acid analysis.

From the foregoing it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.

CLAIMS

1. An isolated polypeptide, comprising at least an immunogenic portion of a breast tumor protein, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(a) sequences recited in SEQ ID NOs: 2-5, 7, 10, 13, 15, 16, 21, 22, 25, 28, 30, 33, 35, 63, 72, 73, 81, 103, 104, 107, 109, 118, 128, 129, 134, 136, 159-161, 164, 181, 183-204 and 210-290;

(b) sequences that hybridize to a sequence recited in any one of SEQ ID NOs: 2-5, 7, 10, 13, 15, 16, 21, 22, 25, 28, 30, 33, 35, 63, 72, 73, 81, 103, 104, 107, 109, 118, 128, 129, 134, 136, 159-161, 164, 181, 183-204 and 210-290 under moderately stringent conditions; and

(c) complements of sequences of (a) or (b).

2. An isolated polypeptide according to claim 1, wherein the polypeptide comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NOs: 2-5, 7, 10, 13, 15, 16, 21, 22, 25, 28, 30, 33, 35, 63, 72, 73, 81, 103, 104, 107, 109, 118, 128, 129, 134, 136, 159-161, 164, 181, 183-204 and 210-290 or a complement of any of the foregoing polynucleotide sequences.

3. An isolated polynucleotide encoding at least 15 amino acid residues of a breast tumor protein, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide comprising a sequence recited in any one of SEQ ID NOs: 2-5, 7, 10, 13, 15, 16, 21, 22,

25, 28, 30, 33, 35, 63, 72, 73, 81, 103, 104, 107, 109, 118, 128, 129, 134, 136, 159-161, 164, 181, 183-204 and 210-290 or a complement of any of the foregoing sequences.

4. An isolated polynucleotide encoding a breast tumor protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide comprising a sequence recited in any one of SEQ ID NOS: 2-5, 7, 10, 13, 15, 16, 21, 22, 25, 28, 30, 33, 35, 63, 72, 73, 81, 103, 104, 107, 109, 118, 128, 129, 134, 136, 159-161, 164, 181, 183-204 and 210-290 or a complement of any of the foregoing sequences.

5. An isolated polynucleotide, comprising a sequence recited in any one of SEQ ID NOS: 2-5, 7, 10, 13, 15, 16, 21, 22, 25, 28, 30, 33, 35, 63, 72, 73, 81, 103, 104, 107, 109, 118, 128, 129, 134, 136, 159-161, 164, 181, 183-204 and 210-290.

6. An isolated polynucleotide, comprising a sequence that hybridizes to a sequence recited in any one of SEQ ID NOS: 2-5, 7, 10, 13, 15, 16, 21, 22, 25, 28, 30, 33, 35, 63, 72, 73, 81, 103, 104, 107, 109, 118, 128, 129, 134, 136, 159-161, 164, 181, 183-204 and 210-290 under moderately stringent conditions.

7. An isolated polynucleotide complementary to a polynucleotide according to any one of claims 3-6.

8. An expression vector, comprising a polynucleotide according to any one of claims 3-7.

9. A host cell transformed or transfected with an expression vector according to claim 8.

10. An isolated antibody, or antigen-binding fragment thereof, that specifically binds to a breast tumor protein that comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NOS: 2-5, 7, 10, 13, 15, 16, 21, 22, 25, 28, 30, 33, 35, 63, 72, 73, 81, 103, 104, 107, 109, 118, 128, 129,

134, 136, 159-161, 164, 181, 183-204 and 210-290 or a complement of any of the foregoing polynucleotide sequences.

11. A fusion protein, comprising at least one polypeptide according to claim 1.

12. A fusion protein according to claim 11, wherein the fusion protein comprises an expression enhancer that increases expression of the fusion protein in a host cell transfected with a polynucleotide encoding the fusion protein.

13. A fusion protein according to claim 11, wherein the fusion protein comprises a T helper epitope that is not present within the polypeptide of claim 1.

14. A fusion protein according to claim 11, wherein the fusion protein comprises an affinity tag.

15. An isolated polynucleotide encoding a fusion protein according to claim 11.

16. A pharmaceutical composition, comprising a physiologically acceptable carrier and at least one component selected from the group consisting of:

- (a) a polypeptide according to claim 1;
- (b) a polynucleotide according to claim 3;
- (c) an antibody according to claim 10;
- (d) a fusion protein according to claim 11; and
- (e) a polynucleotide according to claim 15.

17. A vaccine comprising an immunostimulant and at least one component selected from the group consisting of:

- (a) a polypeptide according to claim 1;

- (b) a polynucleotide according to claim 3;
- (c) an antibody according to claim 10;
- (d) a fusion protein according to claim 11; and
- (e) a polynucleotide according to claim 15.

18. A vaccine according to claim 17, wherein the immunostimulant is an adjuvant.

19. A vaccine according to any claim 17, wherein the immunostimulant induces a predominantly Type I response.

20. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a pharmaceutical composition according to claim 16.

21. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a vaccine according to claim 17.

22. A pharmaceutical composition comprising an antigen-presenting cell that expresses a polypeptide according to claim 1, in combination with a pharmaceutically acceptable carrier or excipient.

23. A pharmaceutical composition according to claim 22, wherein the antigen presenting cell is a dendritic cell or a macrophage.

24. A vaccine comprising an antigen-presenting cell that expresses a polypeptide comprising at least an immunogenic portion of a breast tumor protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

- (a) sequences recited in SEQ ID NOs: 1-38, 42-204, 205, 207 and 210-290;
- (b) sequences that hybridize to a sequence recited in any one of SEQ ID NOs: 1-38, 42-204, 205, 207 and 210-290 under moderately stringent conditions; and
- (c) complements of sequences of (i) or (ii); in combination with an immunostimulant.

25. A vaccine according to claim 24, wherein the immunostimulant is an adjuvant.

26. A vaccine according to claim 24, wherein the immunostimulant induces a predominantly Type I response.

27. A vaccine according to claim 24, wherein the antigen-presenting cell is a dendritic cell.

28. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of an antigen-presenting cell that expresses a polypeptide comprising at least an immunogenic portion of a breast tumor protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

- (a) sequences recited in SEQ ID NOs: 1-38, 42-204, 205, 207 and 210-290;
- (b) sequences that hybridize to a sequence recited in any one of SEQ ID NOs: 1-38, 42-204, 205, 207 and 210-290 under moderately stringent conditions; and
- (c) complements of sequences of (i) or (ii) encoded by a polynucleotide recited in any one of SEQ ID NOs: 1-38, 42-204, 205, 207 and 210-290; and thereby inhibiting the development of a cancer in the patient.

29. A method according to claim 28, wherein the antigen-presenting cell is a dendritic cell.

30. A method according to any one of claims 20, 21 and 28, wherein the cancer is breast cancer.

31. A method for removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a breast tumor protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

- (i) polynucleotides recited in any one of SEQ ID NOS: 1-38, 42-204, 205, 207 and 210-290; and
- (ii) complements of the foregoing polynucleotides;

wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the antigen from the sample.

32. A method according to claim 31, wherein the biological sample is blood or a fraction thereof.

33. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient a biological sample treated according to the method of claim 31.

34. A method for stimulating and/or expanding T cells specific for a breast tumor protein, comprising contacting T cells with at least one component selected from the group consisting of:

- (a) polypeptides comprising at least an immunogenic portion of a breast tumor protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

- (i) sequences recited in SEQ ID NOs: 1-38, 42-204, 205, 207 and 210-290;
- (ii) sequences that hybridize to a sequence recited in any one of SEQ ID NOs: 1-38, 42-204, 205, 207 and 210-290 under moderately stringent conditions; and
- (iii) complements of sequences of (i) or (ii);

- (b) polynucleotides encoding a polypeptide of (a); and
- (c) antigen presenting cells that express a polypeptide of (a); under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells.

35. An isolated T cell population, comprising T cells prepared according to the method of claim 34.

36. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a T cell population according to claim 35.

37. A method for inhibiting the development of a cancer in a patient, comprising the steps of:

(a) incubating CD4⁺ and/or CD8⁺ T cells isolated from a patient with at least one component selected from the group consisting of:

(i) polypeptides comprising at least an immunogenic portion of a breast tumor protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(1) sequences recited in SEQ ID NOs: 1-38, 42-204, 205, 207 and 210-290;

(2) sequences that hybridize to a sequence recited in any one of SEQ ID NOs: 1-38, 42-204, 205, 207 and 210-290 under moderately stringent conditions; and

- (3) complements of sequences of (1) or (2);
- (ii) polynucleotides encoding a polypeptide of (i); and
- (iii) antigen presenting cells that expresses a polypeptide of (i);
 - such that T cells proliferate; and

(b) administering to the patient an effective amount of the proliferated T cells, and thereby inhibiting the development of a cancer in the patient.

38. A method for inhibiting the development of a cancer in a patient, comprising the steps of:

- (a) incubating CD4⁺ and/or CD8⁺ T cells isolated from a patient with at least one component selected from the group consisting of:
 - (i) polypeptides comprising at least an immunogenic portion of a breast tumor protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:
 - (1) sequences recited in SEQ ID NOs: 1-38, 42-204, 205, 207 and 210-290;
 - (2) sequences that hybridize to a sequence recited in any one of SEQ ID NOs: 1-38, 42-204, 205, 207 and 210-290 under moderately stringent conditions; and
 - (3) complements of sequences of (1) or (2);
 - (ii) polynucleotides encoding a polypeptide of (i); and
 - (iii) antigen presenting cells that express a polypeptide of (i);
 - such that T cells proliferate;
- (b) cloning at least one proliferated cell to provide cloned T cells; and
- (c) administering to the patient an effective amount of the cloned T cells, and thereby inhibiting the development of a cancer in the patient.

39. A method for determining the presence or absence of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient with a binding agent that binds to a breast tumor protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NOs: 1-38, 42-204, 205, 207 and 210-290 or a complement of any of the foregoing polynucleotide sequences;

(b) detecting in the sample an amount of polypeptide that binds to the binding agent; and

(c) comparing the amount of polypeptide to a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.

40. A method according to claim 39, wherein the binding agent is an antibody.

41. A method according to claim 40, wherein the antibody is a monoclonal antibody.

42. A method according to claim 40, wherein the cancer is breast cancer.

43. A method for monitoring the progression of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a breast tumor protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NOs: 1-38, 42-204, 205, 207 and 210-290 or a complement of any of the foregoing polynucleotide sequences;

(b) detecting in the sample an amount of polypeptide that binds to the binding agent;

(c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and

(d) comparing the amount of polypeptide detected in step (c) to the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

44. A method according to claim 43, wherein the binding agent is an antibody.

45. A method according to claim 44, wherein the antibody is a monoclonal antibody.

46. A method according to claim 43, wherein the cancer is a breast cancer.

47. A method for determining the presence or absence of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a breast tumor protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 1-38, 42-204, 205, 207 and 210-290 or a complement of any of the foregoing polynucleotide sequences;

(b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; and

(c) comparing the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.

48. A method according to claim 47, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a polymerase chain reaction.

49. A method according to claim 47, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a hybridization assay.

50. A method for monitoring the progression of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a breast tumor protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 1-38, 42-204, 205, 207 and 210-290 or a complement of any of the foregoing polynucleotide sequences;

(b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide;

(c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and

(d) comparing the amount of polynucleotide detected in step (c) to the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

51. A method according to claim 50, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a polymerase chain reaction.

52. A method according to claim 50, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a hybridization assay.

53. A diagnostic kit, comprising:

(a) one or more antibodies according to claim 10; and
(b) a detection reagent comprising a reporter group.

54. A kit according to claim 53, wherein the antibodies are immobilized on a solid support.

55. A kit according to claim 53, wherein the detection reagent comprises an anti-immunoglobulin, protein G, protein A or lectin.

56. A kit according to claim 53, wherein the reporter group is selected from the group consisting of radioisotopes, fluorescent groups, luminescent groups, enzymes, biotin and dye particles.

57. An oligonucleotide comprising 10 to 40 contiguous nucleotides that hybridize under moderately stringent conditions to a polynucleotide that encodes a breast tumor protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NOs: 2-5, 7, 10, 13, 15, 16, 21, 22, 25, 28, 30, 33, 35, 63, 72, 73, 81, 103, 104, 107, 109, 118, 128, 129, 134, 136, 159-161, 164, 181, 183-204 and 210-290 or a complement of any of the foregoing polynucleotides.

58. A oligonucleotide according to claim 57, wherein the oligonucleotide comprises 10-40 contiguous nucleotides recited in any one of SEQ ID NOs: 2-5, 7, 10, 13, 15, 16, 21, 22, 25, 28, 30, 33, 35, 63, 72, 73, 81, 103, 104, 107, 109, 118, 128, 129, 134, 136, 159-161, 164, 181, 183-204 and 210-290.

59. A diagnostic kit, comprising:
(a) an oligonucleotide according to claim 58; and
(b) a diagnostic reagent for use in a polymerase chain reaction or hybridization assay.

SEQUENCE LISTING

<110> Corixa Corporation
Dillon, Davin C.
Day, Craig H.
Jiang, Yuqiu
Wang, Aijun
Houghton, Raymond L.
Mitcham, Jennifer L.

<120> COMPOSITIONS AND METHODS FOR THERAPY AND
DIAGNOSIS OF BREAST CANCER

<130> 210121.491PC

<140> PCT
<141> 2000-11-29

<160> 290

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 298
<212> DNA
<213> Homo sapien

<400> 1
ctgaacagtgc tagctccgt gctggagaca gtcctgctga tcacctgaat gctgaacatg 60
cttcgtgggg ctatctttg tttctctgt agtctcttg gtgatctcat ctgctttct 120
gctcgagtga tgacagcctt .aaccttgc ttccttgc tcagagggga aaaaggaatt 180
ggatttcctc agggctctggg gcctggctg tggcttgagg ttccgagact gatgaatcca 240
agcatgcttgc agggccttgc ccggggcat gcgaaagagaa ggttcccata ccaaacac 298

<210> 2
<211> 276
<212> DNA
<213> Homo sapien

<400> 2
tggaaagggtgt ggtgactaag ggccacgggtt attgggtgaa atttgagatt gtaggccaac 60
tgtattttca agcttctgaa cttaggccaa atattcatcg caaagtctct agcgtcatat 120
ttttcttcacc taaattacgt ttccacgaga ttattttat atagttggtc tatctctgca 180
gtccttgcag gtgaagttgt gtgttactag gctgtgtttt gggatgtcag cagtggcctg 240
aagtgagttgc tgcaataaat gttaagttga aacctc 276

<210> 3
<211> 405
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(405)
<223> n = A,T,C or G

<400> 3

aaatacagt	atagtttgc	tttcttctat	agaatgaaca	tagacataac	cctgaagctt	240
ttagttaca	gggagttcc	atgaagccac	aaactaaact	aattatcaaa	cacattagtt	300
atttccagac	tcaaata	agacatcaa	ccaataaact	gagaagaag	catttcatgt	360
tctcttcat	tttgcataa	agcattttt	ctttgacta	aatgcaagt	gagaattgt	420
atttttctc	ctttaattt	acctcagaag	atgcactatc	taattcatga	gaaatacgaa	480
attcagg	tttatcttct	tccttactt	tgggtctac	aaccagcata	tcttcatggc	540
tgtgaaattc	atggctg					557

<210> 7
 <211> 653
 <212> DNA
 <213> Homo sapien

<400> 7						
cattgtttg	gggaaagt	ggaatattat	tgaggcagg	taagaaatgg	tttacaattc	60
tgaaaggat	atcaaagaaa	aactcattgt	tgagaaagt	atatgagtag	agacctgaaa	120
taagtgg	agtgcgggt	tatgtccagg	gcaataatgt	ttctgacaga	gggagagtc	180
atttcagaag	cctagaggc	tgtgtaa	tgtagaaatg	ccagacagtc	accaggccaa	240
gatgtcgaga	tatccataa	tgaaggggaa	agaaaatac	aatgaaggca	gagaatcac	300
aaaattggat	aagtggc	ttgtaggc	tgtatgttt	agttcatact	aaaattgagt	360
taggctgcca	ttgttagggtt	tgtgagctc	gggataacat	ggtctgaatt	ttat	420
aaggatcact	ccaagtgtt	cattgcaaa	aataacgtaa	ggtggctggt	gttagact	480
aaagtggaat	atagtaacag	tgaatatac	tttggtaa	agcttggtag	atttgaccac	540
acaaaattgt	gaaattac	gtggcacaaa	aaat	ggtacat	aca	600
ccttcgatt	gttttatt	atc	tttttttt	tttttttt	tttttttt	653

<210> 8
 <211> 456
 <212> DNA
 <213> Homo sapien

<400> 8						
cattgtttg	ggcta	ttgtcttat	ccaccctg	tagcaattt	tctaaagct	60
tcaagttc	ccatctaca	tgtgccc	agg	tcaaccaatc	aatggctc	120
caacatgc	cccgccgg	ctg	ggaa	tgtgaa	gttgc	180
gggcgcag	cgaggac	atcc	actgg	gttgc	tttgg	240
gagagac	ccgtg	gaga	ggccg	atcc	ctgtcc	300
taacac	gtgtt	aa	gatc	gttgc	ttggc	360
cagaggag	ggccc	atc	ggat	ttggc	atcat	420
tgat	cgctt	tc	gttgg	atcc	ccgt	456

<210> 9
 <211> 512
 <212> DNA
 <213> Homo sapien

<400> 9						
gtttttgatt	catttttattt	taacaatgtt	taacaatgt	agtccacata	taagatacc	60
aagctt	aatatcata	tataaactg	tttcaacatc	tttggcttca	aaacagtaaa	120
attttttt	caatatcaa	caagt	caa	ttggaaaagg	cataaaatct	180
ctgtatccat	ggagatgtc	tgact	aaatt	cagaat	atct	240
ctttttat	tctgagttc	gtat	ttt	tat	ttttttt	300
tacccctt	aaacagtgtc	tatt	gtt	cat	tttttctat	360
agtcggct	tgg	ttt	ttt	ttt	ttttt	420
aggc	aaat	atc	atc	atc	atc	480
ccac	atc	atc	atc	atc	atc	512

<210> 10

```

<211> 308
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(308)
<223> n = A,T,C or G

<400> 10
atgtttatga agacctttaa atattttat agaaaacaaa tgtcatgca acctaacatc      60
atccattaaa aataaaaagga aaggaaaacg gcagggaaaa gtgcagtaat aacaatggt      120
gacatgcttgc tcttaagca tcatacgaaa ctcatttattt ccaatgaaac aaggatttt      180
agacccatct ttggaaaatga ttcccaaatt aganaaccat caggtctcaa aaaaggaagg      240
gtcatcaaag tccatccagc ccagccaccc tgaggnccct gtatctcctc aacaagccca      300
acacaatg      308

<210> 11
<211> 510
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(510)
<223> n = A,T,C or G

<400> 11
attatatgaa tattttatgt caaaatgctt aacactttaa attagcaaag cgtcattttaa      60
attaaaatttc catttaacta aagatggta accccaanaaa attgtacagt agtgatttc      120
tgctatataa tgccagtctt atgcccataca ataagaactg caacatttgc tgtcacttcc      180
tccatttgc tcttggaccc taaggatga gggaggggac tcagacacaa aacacaaccc      240
aaataaaactg tgcagtgtt cctaatagtt ataaacccaa tctaaggatgt ccaaacagct      300
gaagaataaac tgcagggtt gttccanagc tgatacgagg ttttgccttt acagcctggt      360
aaaagttctg cacttaggtgaa gaagtccacag tttaaggatgt catgttctgt aaatagttac      420
tacatataca cattttactgt ctgttaaacac tagaaatata cattagacag agtaccctca      480
caagttgggt acagtttaaa aaagaagatg      510

<210> 12
<211> 611
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(611)
<223> n = A,T,C or G

<400> 12
agttttataaa aatattttat ttacagttaga gctttacaaa aatagtctta aattaataca      60
aatccctttt gcaatataac ttatatgact atcttctcaa aaacgtgaca ttgcattata      120
acacataaaac tacattttata gttgttaagt cacctttag tataaatatg ttttcatttt      180
ttttttgtaa taaggnacat accaataaca atgaacaatg gacaacaaat cttatgtt      240
tatttttcca atgtaaaattt catctctggc caaaaacaaa ttaaccaaag aaaagtaaaa      300
caattgtccc tctgttcaac aatacagtcc ttttaattt tttgagatgt tatctgacag      360
agacacacagca ttaaactgaa agcaccatgg cataaaagtct agtaacattt tcctcaaaag      420
ctttttccaa tgcctttccct tcaactgttt attcagtatt tggccagtagc aaataaagat      480
tggtctcaac tctcttttc attagtctca agtgttccctt ttagtgcactg agtttcaga      540

```

ccttcccaac tggcatgtgt tttaagtgtg agtttcttc tttggcttca agtggagttt	600
cacaacattt a	611
<210> 13	
<211> 394	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(394)	
<223> n = A,T,C or G	
<400> 13	
caatgttag attcattttt ttagtggcat atacaaagca ccatataata tatgaaacgt	60
anaacaatca tgactatgtt attaactgtt naaataactg ctaanaaaat atagcaatat	120
ttaacacagg atttctaaaa ccattatatt ttcttactt ttcccaaaagc taatgtccca	180
tgttttattt tatanacttt gtttatcaag atttataatgc atttggcacc ttttgggct	240
gaaaatagtt gatgtactct gtacagtaat gtacagttt tataaaaaat tcanaaaat	300
tgcatttggaa atagtcttta tggcttctt ccaagtattt agttcacac aacagcaaac	360
actctgaatg ctttcctcc tgcccaacac aatg	394
<210> 14	
<211> 361	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(361)	
<223> n = A,T,C or G	
<400> 14	
agcagggnact ataattttat aattaattttt acaattcatg tagcaaatgg aaaatcatac	60
agagaggcca atgttatataa ataagagttt atacagaaac tgcccaattca caaaacagca	120
ctgcatggtt tctatattgc aagcacaaga catggtcaca tggttccact gtacaggttag	180
aaacaagccc acagacaata catagagtac cacctgaaac gaggcccttg gagctgctca	240
gcttcttana aaataganaa ctttcaatgg tcataataca ttttgatca aaatgtcttc	300
taaaatgttt tcattgtggg agaaaattaa gaaggggcaa aaatccatct atgaaacttc	360
t	361
<210> 15	
<211> 537	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(537)	
<223> n = A,T,C or G	
<400> 15	
acttacaaaa ttaattttat tttgcaaaac tcaacaaaata cacgttcaga tctggttct	60
cttcaaaaaca tgggtttgtt ttttaacaa acatgcaagt taatttggca tgccaaacat	120
ctttctctt agctcgccctt ggaaaaattt ttttctataac acaaacaagg gtgaaat	180
tgtccaaacc tatttacatt tttaccctctt agaattacat acattaat ttattggag	240
gaaagcaaaa ctgcaaaaaca tagtcttgg cattcacatt tgcttcagca gtataattaa	300
aaccttatat ttgtttaaa gataaacagt ttgaaggaaa ttaataat cttgtttgg	360

ctctgcaaag gagccactat atcaaagcat ttaactggag ctgtttagtt cctgctggta	420
gaatattact tccagccat ttattagctt gtcttccggc ggcccaatac atgtttttt	480
ccctctacac tgaatgaaag tacaaaaaga aaaccatttc tttccccaa cacaatg	537
<210> 16	
<211> 547	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(547)	
<223> n = A,T,C or G	
<400> 16	
gggtgtggng atgtatttat tcataatata tttcagaac acattaataa tggagaataa	60
cacttattca tatactgaat ataactttc ctggagcact cttagagttt tttggagttt	120
gagaataactg ccaggctttt cctaactctt ttgttctttt gaagtggca gggtttctca	180
aaccaagtgt ctccatggg ccattggca aggcttccct tcatcagctt ggagggcag	240
aaagaccatg gcttcagcag ttccatttt gaaaagaagta aaaaaaaagt gaattaatga	300
gcaatcgaa agactcaaag cattttgtac tccacagttc atttcttac acaaacgtcc	360
attactgcag cgggcatgaa aaccggcagg gtgttaggct catggcttga agagaagtca	420
catcaccagc cgatgtttt atgaaaaagg caatcgttcat gattcanaac ctggttctga	480
atttctccag gtgtgcttgcg gagctgaagg tcatgccccat tctgtgcac tctgtgccccaa	540
cacaatg	547
<210> 17	
<211> 342	
<212> DNA	
<213> Homo sapien	
<400> 17	
acattaagaa gctcccttcc tagcatgtcc ttaagaagcc tgcgttgcag cactttcata	60
tcttcttca tcaaacacat ctggatgtt aaaaacagttt cttcaactatc agtattacag	120
aagacacttt tagccatgtt agttttcaaa agaagaaaagc ctctgttgcg cgttttttgc	180
atatgcactg aacttctgaa atatctttc cccaaagtcc acaaatttctt tttccaaatc	240
ttttaagac tgcgttgcg tttcaaaattt ctccagctcc tctatgataa tgaattggaa	300
tttatcaagt ttttaatcc tagagtcgttgcg actttggatg at	342
<210> 18	
<211> 279	
<212> DNA	
<213> Homo sapien	
<400> 18	
catcataagg ttttattcat atatatacag ggtattaaga attaagagga tgctggctc	60
tgttcttggc ttggaaagatt ctatttaattt gaaactctt gttcagaaag caataacttt	120
gtctcggtcc tggtggctt aaccctaagg tgagtgttgcg gtacagtgtg tgcgttgcg	180
atggagattt ggaatttgcac tctctgcctt taaaatgttcc cccaaataattt gttgtgttgc	240
tgatacgttgcg ataataaaag tattcttgcg agaatctgtt	279
<210> 19	
<211> 239	
<212> DNA	
<213> Homo sapien	
<400> 19	
ctgcacggcg ttttgcgttgcg ctgcgttgcg cctggccca gtcacgggc agtgggttgc	60

cctaactgcc caggcaggcg agagctactt ccagagcctt ccagtgcattt ggagggcagg
 gcttaggtgta gcgggtgtctc ctctttgaaa ttaagaacta tctttcttgc agcaaagctg
 cacctgatga tgctgcctct cctctctgtg ttgtctggc cttgtttac aagcacgctg
 120
 180
 239

<210> 20
 <211> 527
 <212> DNA
 <213> Homo sapien

<400> 20
 ctgaaccatt atggataaa ctggtgcaaa ttcttgctt tctctacttc tcactgattt
 aacataagct tccagggtctc ccctgatgag gaggagcctg tccttttcag atggatggtc
 60
 120
 180
 240
 300
 360
 420
 480
 527

atccagccac tgagagaagc gtgtgtggg ccactctgctt ctctggaaag gagatttcag
 ttccgggt gctctcgta aaaaaactg aataatgatg ctgaacggaa tcacatcccc
 caatgcagga ctactggcta catgttcaact tgccttggaaag acaagaggctc tgaatgatct
 cagcatccga taggactttc ctaaatcaga tactcgctca cagaatgaac ccacagccaa
 ctccatctgt gcaaaaatcg cagcaagtcg cattttccca ctttcacccaa gaggtcttat
 gagactggca tggcgatcaa aaagttcaac agcttttgg gcaataaccc tcaatgttgc
 aaagacaaaaa tccaaggcatt caaatgtttt aaaaatgtca ctcataaa

<210> 21
 <211> 399
 <212> DNA
 <213> Homo sapien

<400> 21
 ctgcaatgt tgcaagtgtt atttccaccc agctctgact ctccacttctt aaccagacaa
 acagccaaacc aaccatcaa catgttatttta ataaccaccc atgggggtgca aagcacaaaa
 60
 120
 180
 240
 300
 360
 399

gggcactcat cttggaaaagg aaagaccaag aatgtgttagt agtaaagaga cagagaccag
 accctactctt caagatcaag agacttcgtt ctcggagaca tctggcattt ctctttttt
 aataaacccctt atttccctt taaaatacat ttgtttttggg ggcccaagaat caagaaagga
 aactttacaa agtaaacaga agttactccc cacagggagg cagaagcaga ttaaccccaa
 cagcagacat ctgccccggaa gagcaaaactc cacatctgg

<210> 22
 <211> 532
 <212> DNA
 <213> Homo sapien

<400> 22
 ccagaagggt aaaaaaggat atctgataat gctcaaagtg cagtagaaat acttttaacc
 attgtgata caaagagagc tggaaatgaaa gagctaaaac gtcatcctt cttcagtgtat
 60
 120
 180
 240
 300
 360
 420
 480
 532

gtggacttggg aaaatctgca gcatcagact atgccttca tccccccagcc agatgtatgaa
 acagataacctt cctatccatgaa agccaggaat actgctcgtt acctgaccgtt atctggattt
 agtctgttagc aaaaaatttt tccttttagt ctggcctgtt gttatagaat gaaacttgcatt
 aattatataac tccttaataac tagattgtatc taagggggaa agatcattat ttaacccatgt
 tcaatgtgtt ttaatgttac gttacagctt tcacagagtt aaaaggctga aaggaatata
 gtcagtaatt tatcttaacc tcaaaactgtt atataatct tcaaaagctt tttcatctat
 ttatgggtt tattgcattt tatggaaaactt gaagcatcaa taaaattttaga gg

<210> 23
 <211> 215
 <212> DNA
 <213> Homo sapien

<400> 23
 tgcaaataag ggctgctgtt tcgacgacac cgttcggtt gttccctggg gttttatcc
 taataccatc gacgtccctc cagaagagga gtgtgtatcc tagacacttc tgcagggtt
 60
 120

tgccctgcac	ctgacacacgt	gccgtcccc	gcacgggtat	tagtcccaga	gctcggtgc	180
cacccatccacc	ggacacaccca	gacacgttc	tgcag			215
<210> 24						
<211> 215						
<212> DNA						
<213> Homo sapien						
<400> 24						
cctgaggctc	caggctaaga	agttagccaag	tttcacctgg	agagaagagt	agagggactt	60
cccaaatttc	ttcctgaact	cagctctgtat	actcagaagg	tcagtctcac	atcgagagat	120
aaggatgcga	atcaggactt	ggtaattggg	ctcagttcc	tagtagggg	agaaagagat	180
ggggggtagt	tagttagact	ctcactgaga	gtagg			215
<210> 25						
<211> 530						
<212> DNA						
<213> Homo sapien						
<400> 25						
ttttttttct	agtaagacta	gatttattca	ataccctagt	aaaagtttg	attataagta	60
tccaaacagta	taaaaagtac	aaaacagatc	tgttagatttc	taatataat	atacaaagt	120
catgactaca	tacagtagat	cctacaggca	aagagaggtg	gaaggggaaa	aagaagactg	180
tgggtgaggt	ctagtaataa	ataaaataat	acagaagtag	agatgatcca	tattatagta	240
tattctacca	ccaataactgc	agccaaaatg	tacaaaaaaaaa	atcatttcaa	ataactcagg	300
aggatgataa	tggctggact	tttgcatttc	acctcaaaaga	ctgtgggaga	gccaactcaa	360
ctcactgtat	agtctgtgca	tatggggct	tgttagcatgt	aggtttttc	aaaaagaagg	420
aaatataaaaa	tgttttagatt	aagaactata	aaactacagg	gtgcctataa	aaggtggctt	480
actccttatt	gttattatac	tatccaattt	ttaaaatgca	gtttaaaaaa		530
<210> 26						
<211> 366						
<212> DNA						
<213> Homo sapien						
<400> 26						
ccagcagttc	tcggacactcc	tctggggca	gggagaggcc	attgggtcag	gggctggacc	60
caggagggt	tggaatgggt	gaaagatggg	gagcaagttt	ttagggtaca	gggtgggcct	120
aagatgggtc	agtagacaga	tgggagcaca	gagcaggcga	gggggtgagg	tcaagtggagg	180
gccacaggat	gtgctgaggg	ctcccaggga	gccttaccca	ggctcacgtc	ctcctggtca	240
ccacctgtac	tgtctgggt	ccacagggtg	tggcgttgc	caggagcac	tgggaggggcc	300
tgcgttaggt	ccacctgtag	ggagaggatg	tcaggaccac	tagcctctgg	gcaagggcag	360
aggagg						366
<210> 27						
<211> 331						
<212> DNA						
<213> Homo sapien						
<220>						
<221> misc_feature						
<222> (1)...(331)						
<223> n = A,T,C or G						
<400> 27						
ccaaactcag	agatggtacc	agccaggggc	aagcatgacc	agagccaggg	accctgtggc	60
tctgatcccc	catttatcca	ccccatgtgc	ctcaggacta	gagtgagcaa	tcataccctta	120
taaatgactt	ttgtgcctt	ctgctccagt	ctcaaaattt	cctacacccgt	ccagttcttt	180

acattttcc aaggaaaagga aaacggaagc agggttcttg cctggtagct ccaggaccca	240
nctctgcagg cacccaaaga ccctctgtgt ccagccttcc ccttgagttc tcgaaacctc	300
ctccctaatt ctcccttccat tccccacacaag g	331
<210> 28	
<211> 530	
<212> DNA	
<213> Homo sapien	
<400> 28	
ccatgaatgc ccaacaagat aatattctat accagactgt tacaggattg aagaaaagatt	60
tgtcaggagt tcagaagggtc cctgcactcc tagaaaaatca agtggaggaa aggacttgg	120
ctgattcaga agatatttggg agctctgagt gctctgacac agattctgaa gaggcaggag	180
accatgcccgg ccccaagaaa cacaccacgg accctgacat tgataaaaaaa gaaagaaaaaa	240
agatggtaaa ggaagccccag agagagaaaa gaaaaaaacaa aattcttaaa catgtaaaaaa	300
aaagaaaagga gaagacagcc aagacgaaaa aaggcaataa gaatgagaac catattatgt	360
acagtcattt tcctcagttc ctttctcgc ctgaacttta aagctgcatac tggaaagatgg	420
cttattgggtt ttaaccagat tgcatacgat gcactgtctg tgaagacgga ttcaaattgtt	480
ttcatgtaac tatgtaaaaaa gctctaagct cttagtcta gatccagtc	530
<210> 29	
<211> 571	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(571)	
<223> n = A,T,C or G	
<400> 29	
ccataatatt ctgatgatca aggagcacac atataaaaaa gttattggat tactgcaatt	60
ctcagaggca caaaacctga catgggtgtat tatagtatata aatcagtcac gggggggaaa	120
agaacattaa gtctttaaaa aggcttagga agacataaac agtaaatctt tgttttctta	180
ccttccttgc gacagtgttata tatttcaattt tcttcttgc aaaatgttca caaattcatt	240
tgctcaggat ttatataaga taataactta aaacaactaa cagttgttta tgctatatgc	300
atatcatgca tgttctactg gttcaaggac aaaataaaaa caagatctt tctgtaaagc	360
aaatataattt attatgcact ttcataataca cagggatttt ttgagtacca angggataaa	420
ataaaaactt tacaatgtga aattcaatgt acatttttgg ctatttacat acctcaaacc	480
aaggaaaaaaa taaaagaaaaa gcatttggtt gcaactacat ttgctgagaa gtgtaaatgg	540
aggacattaa gcaaaacaaa tatttgcata g	571
<210> 30	
<211> 917	
<212> DNA	
<213> Homo sapien	
<400> 30	
actgccagag agtatgattt gaaggagatg ggagcagatg taattcttgg ctggaatctc	60
tcatttcaaa atcacttcac ataatgggtgt catcattaa acacttaaca gtcagtgcac	120
ctgccactgt aacatctgtt tggacaaaac cacaaggagg gggaggagaa aatgccatca	180
ctattatgtt aacaaacattt taattttaat ggttgcgtca ctatgttattt tctgcagaaaa	240
acagtttac ccccccctt tcacaggatcc aaattaaatca aggtgcattt tctataatct	300
gatgcatttc aaatttagctc atgattcaaa tttgccttc ttgaaagcaca tatacctttt	360
attttaaaag tccattatag agaatttggg atatataagg tatttgcattt gcagaacacc	420
cctctaattt tgcattatata gcaaaacaaa aacagtatca tatacatca gatcatactt	480
ttaaaagtaag tttaaagggtc tcaatttgcac agatattaaa tttatattttt ctttcttattt	540
aaaaatatttattt catttcaattt ttgtatattt gtaacatattt ttaagatgac cagcaagacc	600

tagtcaattt	gaaaataccc	ttgcattcca	tacacaagct	ataccataag	taataaccca	660
agtatatgtat	gtgtaaaagt	tggtgaaggt	cataatactg	aatttttttgc	caaatgtaaa	720
ctgctttcca	agtaatcagc	accattttt	actagactac	attttaatca	cttccttagc	780
tgcctacaac	ctctacttag	gcataaataa	aagaatctga	aattggata	tttccccttc	840
ctgctgtgtt	aaccaaaaat	actatttgac	ttaaagatca	aagagtctt	ttcctgaagg	900
tttttgtttt	taaatgt					917
<210> 31						
<211> 367						
<212> DNA						
<213> Homo sapien						
<220>						
<221> misc_feature						
<222> (1)...(367)						
<223> n = A,T,C or G						
<400> 31						
tcttttcttt	ctgtatttcc	caaattacag	ggagctatgc	ccttggtatt	gcacacagta	60
cactgcaaaa	gattcacaag	gttagtcaa	agtcattttt	gccctggta	ttcaaagctc	120
aaanaatttt	ctagcataaa	gtcttattaa	aaatttttaat	caaaatatta	tttgagttt	180
agtttaataa	aacaataccca	ctatatatac	tctcaacaac	ttcattatata	aatcagtctt	240
atgaggttgt	acttgctttt	catatcacac	tgattaagga	caaaaataat	tttgatgtac	300
atgtaccata	cactgatatg	caatctacac	actgatgcat	ttacatacat	acaaccccaa	360
cacaatg						367
<210> 32						
<211> 847						
<212> DNA						
<213> Homo sapien						
<400> 32						
catttgttgc	ggctggcagg	atagaagcag	cggctcaatt	ggacttttc	accagggaaa	60
tcagagacaa	tgatggggct	cttcccaga	actacagggg	ctctggccat	ttctgtggta	120
agtcctggat	tttcctaata	atcacaaact	tccctgcttc	ctccctgtt	aaagaatatt	180
atatttgatt	gcacaatctt	tattataaaat	tctaaaagga	gtgcagtgg	aatcaacact	240
ttgaaatgaa	atcgtgaaga	ttaccaattt	ccttctttt	ttgtttttt	tgttgtat	300
tacatagaaa	aataaaccag	aaagaaatga	gtttaaaaaa	ccattnagaa	tttttttag	360
ttaatgaatt	aagtaatctt	aatcacaggt	tatattttcc	acaacattt	cacttctt	420
aaagttatgc	ttttacttagt	ttttctaacc	cacaaacaag	aacacaggag	ccacttctat	480
tttccaagat	tacatgtctc	tttagcatata	gctaagaact	ctacacgcct	gggcttgata	540
cctgacacgc	ttttaaaagt	aaaaaatcgc	agaattaaaa	tcaaagcagt	gtttactct	600
agagaagttg	ggaggattat	taagtaagta	tttatgttta	gctattatgt	gccaaaagaa	660
aatgtcagcc	tttggggatg	ggggaaaga	catacaacat	tttaaagcca	ttttttcag	720
aaaagtaata	cttctgttga	ttgagaaagt	cgtacatagt	attatctaaa	agagaaaacgg	780
aatgttacag	actgtttaaa	acctggatgt	tacagactaa	cttactccct	aactgtgttc	840
ttatagc						847
<210> 33						
<211> 863						
<212> DNA						
<213> Homo sapien						
<220>						
<221> misc_feature						
<222> (1)...(863)						
<223> n = A,T,C or G						

<400> 33

cattgtgttgcgttttatt	tgagtttatg aacagaaata gaaagtatgg	tgcttgggtt	60
ttgccttcttactcctga	aagttaaatac agaagacact gatttcattt	tgtgaaattt	120
agctcagaga ctattgatct	tttgtttcat taatatgaac aactattagt	aaaaaatagc	180
tttaacagca ttctctgtga	tatctagtaa tctattctt	taatgtgaaa	240
atgtccttggatcta	gcttaaattt gccagttt	ctgtatgtca	300
ttcccttaag gttggtaata	naattttgtt aatcttttgc	tacctgtatgg	360
aatgtcgatt gggtaattt	aaatttctgtg ctaattttaa	acttaatttgc	420
tgattgtcttctgtatgtt	attgttagtt aatgaagtat	agcttgcac	480
catgggtctgtatgtatgtt	taagtaaaaaa taaataaacc	aatatttgc	540
tttctatcaa accaattttt	ctntaaagac	aatttcactt	600
nttgaatgcc tcaaattaa	cgttatctat	ttaatcttcc	660
aaggagggtg tgatataattt	agggtgtaaat atatcacata	tatggtgtga	720
atttataat tctgtttagt	tttctactct	gactaaaattt	780
taaactaatt tctgtttagt	tttctactct	gcttaccatg	840
taataaaatgt atcagtcnct	agc	tatgcctttt	863

<210> 34

<211> 432

<212> DNA

<213> Homo sapien

<400> 34

agtgcatttc ctcttgcattt	gtctgggtta aaaccattcc	ttttgtatga	aatgttttga	60
cttaggaatc attttatgtt	cttgccttac ctggattgtc	aacaactgaa	agtacatatt	120
tcatccaaat caagctaaaa	tgtatattaag ttgattctga	gagtcaggt	cagtaagcct	180
cattatttgg aatttgcag	aaggatataagg tgatcgatc	tgtttcattt	ataaaaaggc	240
cagtttttag gactagtaca	ttcctgttat tttctgggtt	ttatcatttt	gcctaaaata	300
ggatataaaaa gggacaaaaaa	ataagtagac tttttttatg	tgtgaatttat	atttctacta	360
aatgtttttg tatgactgtt	ttatacttga taatataat	atataatataat	atataatatac	420
acttgcattttttaa	tt			432

<210> 35

<211> 350

<212> DNA

<213> Homo sapien

<400> 35

ccagaggggt gtttatctta	gggttggaaat gtttctgatt	atgctgacaa	tagccattag	60	
gctgtatgtttt	tggggcttgg	tttaggcagt	ttttaaataa	aagagaactt	120
tgtttgtccaa	agatgggtat	gttccctgtc	tcaatttagca	taaacaaaa	180
taccctgttg	aatgttcctc	atccctctga	gcttctccac	tcacaggata	240
tggctcccc	tcatggacac	ctgcaaatgc	agagtgtggg	ggctctctgc	300
actagcaaga	gcaaaaagctg	ctccgagtc	tgtttttttaga	acctgggtcga	350

<210> 36

<211> 1082

<212> DNA

<213> Homo sapien

<400> 36

atgaactaca gcctccactt	ggccttcgtg tgcgttgc	tcttcactga	gaggatgtgc	60	
atccaggggatc	gtcagttcaa cgtcgaggc	ggcagaagt	acaagcttc	120	
tttggaaacc	tccacagcagg	atataacaaa	tttctcaggc	ccaattttgg	180
gtacagatag	cgctgactct	ggacatgc	agtatctcta	gcatttcaga	240
gactacacag	ccaccatata	cctccgacag	cgctggatgg	accagcgct	300
ggcaacaaga	gttcaactct	ggatgcccgc	ctcggtggat	tccctctgggt	360
tacattgtgg	agtccaagaa	gtccttcctc	catgaagtca	ctgtggaaa	420

cgccctttct ccaatggcac ggtcctgtat gcccctcagaa tcacgacaac tggatgtgc	480
aacatggatc tgtctaaata ccccatggac acacagacat gcaaggatgc gctggaaagc	540
tggggctatg atggaaatga tgtggagttc acctggctga gagggaacga ctctgtgcgt	600
ggactggAAC acctgcggct tgctcagttc accatagagc ggtatttcac cttagtcacc	660
agatcgcagc aggagacagg aaattacact agattggct tacagttga gttcggagg	720
aatgttctgt atttcatttt ggtatctctc cgattcagtc cctgcaagaa cctgcattgg	780
ggacaacaaa ggaagtagaa gaagtcagta ttactaatat catcaacagc tccatctcca	840
gctttaaacg gaagatcagc tttgccagca ttgaaatttc cagcgacaaac gttgactaca	900
gtgacttgac aatgaaaacc agcgacaaatg taaagttgtt cttccgagaa aagatggca	960
ggattgtga ttatccaca attccaaacc ccagtaatgt tgatcaat tccaaactac	1020
tgtttctttt gatTTTatg ctggccatg tattttactg ggcataactac atgtatTTT	1080
ga	1082

<210> 37
 <211> 1135
 <212> DNA
 <213> Homo sapien

<400> 37	
atgaactaca gcctccactt ggccttcgtg tgtctgagtc tcttcactga gaggatgtgc	60
atccaggggga gtcagttcaa cgctcgaggc ggcagaagtg acaagcttc cctgcctggc	120
tttggagaacc tcacagcagg atataacaaa ttctcaggc ccaattttgg tggagaaccc	180
gtacagatag cgctgactct ggacatttgc agtatctcta gcatttcaga gagtaacatg	240
gactacacag ccaccatata cctccgacag cgctggatgg accagcgct ggttttggaa	300
ggcaacaaga gtttactct ggatgcccgc ctctggatgt tcctctgggt gccagatact	360
tacattgtgg agtccaaagaa gtccttcctc catgaagtca ctgtggggaaa cagctcattc	420
cgcctttct ccaatggcac ggtcctgtat gcccctcagaa tcacgacaac tggatgtgc	480
aacatggatc tgtctaaata ccccatggac acacagacat gcaaggatgc gctggaaagc	540
tggggctatg atggaaatga tgtggagttc acctggctga gagggaacga ctctgtgcgt	600
ggactggAAC acctgcggct tgctcagttc accatagagc ggtatttcac cttagtcacc	660
agatcgcagc aggagacagg aaattacact agattggct tacagttga gttcggagg	720
aatgttctgt atttcatttt ggtatctctc cgattcagtc cctgcaagaa cctgcattgg	780
ttttttcat ttggatctc ttcgttca gtccttgcac gaaaccgcatt tggggacaac	840
aaaggaagatc gaagaagtcatc gtattactaa tatcatcaac agtccatct ccagctttaa	900
acggaaatgc agcttgcac gcatttgcac ttccagcgac aacgttgcact acagtgcatt	960
gacaatgaaa accagcgaca agttaaagtt tgccttcgc gaaaagatgg gcaggattgt	1020
tgattatttc acaattcaaa accccagttaa ttgttgcac tattccaaac tactgtttcc	1080
tttgatTTTt atgttagccatc atgtatTTTt ctggccatcc tacatgtatt tttggaa	1135

<210> 38
 <211> 1323
 <212> DNA
 <213> Homo sapien

<400> 38	
atgaactaca gcctccactt ggccttcgtg tgtctgagtc tcttcactga gaggatgtgc	60
atccaggggga gtcagttcaa cgctcgaggc ggcagaagtg acaagcttc cctgcctggc	120
tttggagaacc tcacagcagg atataacaaa ttctcaggc ccaattttgg tggagaaccc	180
gtacagatag cgctgactct ggacatttgc agtatctcta gcatttcaga gagtaacatg	240
gactacacag ccaccatata cctccgacag cgctggatgg accagcgct ggttttggaa	300
ggcaacaaga gtttactct ggatgcccgc ctctggatgt tcctctgggt gccagatact	360
tacattgtgg agtccaaagaa gtccttcctc catgaagtca ctgtggggaaa cagctcattc	420
cgcctttct ccaatggcac ggtcctgtat gcccctcagaa tcacgacaac tggatgtgc	480
aacatggatc tgtctaaata ccccatggac acacagacat gcaaggatgc gctggaaagc	540
tggggctatg atggaaatga tgtggagttc acctggctga gagggaacga ctctgtgcgt	600
ggactggAAC acctgcggct tgctcagttc accatagagc ggtatttcac cttagtcacc	660
agatcgcagc aggagacagg aaattacact agattggct tacagttga gttcggagg	720
aatgttctgt atttcatttt ggtatctctc cgattcagtc cctgcaagaa cctgcattgg	780

tgggtttcat tttggatctc tctcgattca gtcctgcaa gaacctgcat tggagtgacg	840
accgtgttat caatgaccac actgatgatc gggcccgca cttctttcc caacaccaac	900
tgcttcatca aggccatcg tttgtacctg gggatctgtt ttagctttgt gttggggcc	960
ttgctagaat atgcgttgc tcactacatg tccctacagc agatggcagc caaagatagg	1020
gggacaacaa aggaagttaga agaagttagt attactaata tcatcaacag ctccatctcc	1080
agctttaaac ggaagatcgat ctttgcagc attgaaattt ccagcgacaa cgttgactac	1140
agtgacttga caatgaaaac cagcgacaaag ttcaagttt tcttccgaga aaagatgggc	1200
aggattgtt attatttcac aattcaaaac cccagtaatg ttgatacta ttccaaacta	1260
ctgtttcctt tgattttat gtagccat gtatTTTact gggcatacta catgtatTTT	1320
tga	1323

<210> 39
 <211> 440
 <212> PRT
 <213> Homo sapien

<400> 39															
Met	Asn	Tyr	Ser	Leu	His	Leu	Ala	Phe	Val	Cys	Leu	Ser	Leu	Phe	Thr
1				5					10					15	
Glu	Arg	Met	Cys	Ile	Gln	Gly	Ser	Gln	Phe	Asn	Val	Glu	Val	Gly	Arg
		20						25				30			
Ser	Asp	Lys	Leu	Ser	Leu	Pro	Gly	Phe	Glu	Asn	Leu	Thr	Ala	Gly	Tyr
	35					40						45			
Asn	Lys	Phe	Leu	Arg	Pro	Asn	Phe	Gly	Gly	Glu	Pro	Val	Gln	Ile	Ala
	50					55				60					
Leu	Thr	Leu	Asp	Ile	Ala	Ser	Ile	Ser	Ser	Ile	Ser	Glu	Ser	Asn	Met
	65					70				75				80	
Asp	Tyr	Thr	Ala	Thr	Ile	Tyr	Leu	Arg	Gln	Arg	Trp	Met	Asp	Gln	Arg
		85						90				95			
Leu	Val	Phe	Glu	Gly	Asn	Lys	Ser	Phe	Thr	Leu	Asp	Ala	Arg	Leu	Val
	100					105						110			
Glu	Phe	Leu	Trp	Val	Pro	Asp	Thr	Tyr	Ile	Val	Glu	Ser	Lys	Lys	Ser
	115					120						125			
Phe	Leu	His	Glu	Val	Thr	Val	Gly	Asn	Arg	Leu	Ile	Arg	Leu	Phe	Ser
	130					135						140			
Asn	Gly	Thr	Val	Leu	Tyr	Ala	Leu	Arg	Ile	Thr	Thr	Thr	Val	Ala	Cys
	145					150				155				160	
Asn	Met	Asp	Leu	Ser	Lys	Tyr	Pro	Met	Asp	Thr	Gln	Thr	Cys	Lys	Leu
		165						170				175			
Gln	Leu	Glu	Ser	Trp	Gly	Tyr	Asp	Gly	Asn	Asp	Val	Glu	Phe	Thr	Trp
	180					185						190			
Leu	Arg	Gly	Asn	Asp	Ser	Val	Arg	Gly	Leu	Glu	His	Leu	Arg	Leu	Ala
	195					200						205			
Gln	Tyr	Thr	Ile	Glu	Arg	Tyr	Phe	Thr	Leu	Val	Thr	Arg	Ser	Gln	Gln
	210					215						220			
Glu	Thr	Gly	Asn	Tyr	Thr	Arg	Leu	Val	Leu	Gln	Phe	Glu	Leu	Arg	Arg
	225					230				235				240	
Asn	Val	Leu	Tyr	Phe	Ile	Leu	Glu	Thr	Tyr	Val	Pro	Ser	Thr	Phe	Leu
		245						250				255			
Val	Val	Leu	Ser	Trp	Val	Ser	Phe	Trp	Ile	Ser	Leu	Asp	Ser	Val	Pro
		260					265					270			
Ala	Arg	Thr	Cys	Ile	Gly	Val	Thr	Thr	Val	Leu	Ser	Met	Thr	Thr	Leu
	275					280						285			
Met	Ile	Gly	Ser	Arg	Thr	Ser	Leu	Pro	Asn	Thr	Asn	Cys	Phe	Ile	Lys
	290					295						300			
Ala	Ile	Asp	Val	Tyr	Leu	Gly	Ile	Cys	Phe	Ser	Phe	Val	Phe	Gly	Ala
	305					310				315				320	
Leu	Leu	Glu	Tyr	Ala	Val	Ala	His	Tyr	Ser	Ser	Leu	Gln	Gln	Met	Ala

325	330	335
Ala Lys Asp Arg Gly Thr Thr Lys Glu Val Glu Glu Val Ser Ile Thr		
340	345	350
Asn Ile Ile Asn Ser Ser Ile Ser Ser Phe Lys Arg Lys Ile Ser Phe		
355	360	365
Ala Ser Ile Glu Ile Ser Ser Asp Asn Val Asp Tyr Ser Asp Leu Thr		
370	375	380
Met Lys Thr Ser Asp Lys Phe Lys Phe Val Phe Arg Glu Lys Met Gly		
385	390	395
Arg Ile Val Asp Tyr Phe Thr Ile Gln Asn Pro Ser Asn Val Asp His		400
405	410	415
Tyr Ser Lys Leu Leu Phe Pro Leu Ile Phe Met Leu Ala Asn Val Phe		
420	425	430
Tyr Trp Ala Tyr Tyr Met Tyr Phe		
435	440	

<210> 40
 <211> 289
 <212> PRT
 <213> Homo sapien

<400> 40			
Met Asn Tyr Ser Leu His Leu Ala Phe Val Cys Leu Ser Leu Phe Thr			
1	5	10	15
Glu Arg Met Cys Ile Gln Gly Ser Gln Phe Asn Val Glu Val Gly Arg			
20	25	30	
Ser Asp Lys Leu Ser Leu Pro Gly Phe Glu Asn Leu Thr Ala Gly Tyr			
35	40	45	
Asn Lys Phe Leu Arg Pro Asn Phe Gly Gly Glu Pro Val Gln Ile Ala			
50	55	60	
Leu Thr Leu Asp Ile Ala Ser Ile Ser Ser Ile Ser Glu Ser Asn Met			
65	70	75	80
Asp Tyr Thr Ala Thr Ile Tyr Leu Arg Gln Arg Trp Met Asp Gln Arg			
85	90	95	
Leu Val Phe Glu Gly Asn Lys Ser Phe Thr Leu Asp Ala Arg Leu Val			
100	105	110	
Glu Phe Leu Trp Val Pro Asp Thr Tyr Ile Val Glu Ser Lys Lys Ser			
115	120	125	
Phe Leu His Glu Val Thr Val Gly Asn Arg Leu Ile Arg Leu Phe Ser			
130	135	140	
Asn Gly Thr Val Leu Tyr Ala Leu Arg Ile Thr Thr Thr Val Ala Cys			
145	150	155	160
Asn Met Asp Leu Ser Lys Tyr Pro Met Asp Thr Gln Thr Cys Lys Leu			
165	170	175	
Gln Leu Glu Ser Trp Gly Tyr Asp Gly Asn Asp Val Glu Phe Thr Trp			
180	185	190	
Leu Arg Gly Asn Asp Ser Val Arg Gly Leu Glu His Leu Arg Leu Ala			
195	200	205	
Gln Tyr Thr Ile Glu Arg Tyr Phe Thr Leu Val Thr Arg Ser Gln Gln			
210	215	220	
Glu Thr Gly Asn Tyr Thr Arg Leu Val Leu Gln Phe Glu Leu Arg Arg			
225	230	235	240
Asn Val Leu Tyr Phe Ile Leu Glu Thr Tyr Val Pro Ser Thr Phe Leu			
245	250	255	
Val Val Leu Ser Trp Val Ser Phe Trp Ile Ser Leu Asp Ser Val Pro			
260	265	270	
Ala Arg Thr Arg Ile Gly Asp Asn Lys Gly Ser Arg Arg Ser Gln Tyr			
275	280	285	

Tyr

<210> 41
 <211> 265
 <212> PRT
 <213> Homo sapien

<400> 41
 Met Asn Tyr Ser Leu His Leu Ala Phe Val Cys Leu Ser Leu Phe Thr
 1 5 10 15
 Glu Arg Met Cys Ile Gln Gly Ser Gln Phe Asn Val Glu Val Gly Arg
 20 25 30
 Ser Asp Lys Leu Ser Leu Pro Gly Phe Glu Asn Leu Thr Ala Gly Tyr
 35 40 45
 Asn Lys Phe Leu Arg Pro Asn Phe Gly Gly Glu Pro Val Gln Ile Ala
 50 55 60
 Leu Thr Leu Asp Ile Ala Ser Ile Ser Ser Ile Ser Glu Ser Asn Met
 65 70 75 80
 Asp Tyr Thr Ala Thr Ile Tyr Leu Arg Gln Arg Trp Met Asp Gln Arg
 85 90 95
 Leu Val Phe Glu Gly Asn Lys Ser Phe Thr Leu Asp Ala Arg Leu Val
 100 105 110
 Glu Phe Leu Trp Val Pro Asp Thr Tyr Ile Val Glu Ser Lys Lys Ser
 115 120 125
 Phe Leu His Glu Val Thr Val Gly Asn Arg Leu Ile Arg Leu Phe Ser
 130 135 140
 Asn Gly Thr Val Leu Tyr Ala Leu Arg Ile Thr Thr Thr Val Ala Cys
 145 150 155 160
 Asn Met Asp Leu Ser Lys Tyr Pro Met Asp Thr Gln Thr Cys Lys Leu
 165 170 175
 Gln Leu Glu Ser Trp Gly Tyr Asp Gly Asn Asp Val Glu Phe Thr Trp
 180 185 190
 Leu Arg Gly Asn Asp Ser Val Arg Gly Leu Glu His Leu Arg Leu Ala
 195 200 205
 Gln Tyr Thr Ile Glu Arg Tyr Phe Thr Leu Val Thr Arg Ser Gln Gln
 210 215 220
 Glu Thr Gly Asn Tyr Thr Arg Leu Val Leu Gln Phe Glu Leu Arg Arg
 225 230 235 240
 Asn Val Leu Tyr Phe Ile Leu Asp Leu Ser Arg Phe Ser Pro Cys Lys
 245 250 255
 Asn Leu His Trp Gly Gln Gln Arg Lys
 260 265

<210> 42
 <211> 574
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(574)
 <223> n = A,T,C or G

<400> 42
 accaacanag ctttagtaatt tctaaaaaga aaaaatgatc tttttccgac ttctaaacaa 60
 gtgactatac tagcataaat cattcttcta gtaaaacagc taaggtatag acattctaat 120
 aatttggaa aacctatgtatacgtaaa aactcagaaa tgcaaagatg ttggttttt 180

gtttctca	gt	ctgttttagc	ttttaactct	gaaaacgc	at	gcacactgaa	ctctgctca	240	
tgctaaacag	tg	tcaccagcag	gttcc	cagg	ttt	cagccc	taaaatgtaa	aacctggata	300
atcagtgtat	at	tg	tcaccag	aatcagc	catt	ttttttta	ctgcaaaaaa	tgatggtctc	360
atctctgaat	at	tc	at	cattttt	ttt	tttgc	aacataactat	agctaata	420
ctaaattgct	ct	ta	at	ttatattt	ttt	tttgc	atgtaaaaca	aaagataat	480
ataggaaaaa	at	tt	at	ttatattt	ttt	tttgc	actttcgat	aaagtaaaatt	540
tattcatttt	tat	tt	at	ttatattt	ttt	tttgc	tttgcattat	tttgcattat	574
<210> 43									
<211> 467									
<212> DNA									
<213> Homo sapien									
<220>									
<221> misc_feature									
<222> (1)...(467)									
<223> n = A,T,C or G									
<400> 43									
tttttttttt	ttttttttt	tttttttattt	ccatcaattt	ataaaataa	acatgtatag	cagg	tttca	60	
caattgtctt	ca	attgtt	gtt	aaaaagac	ataagaaaga	gaagg	gttgg	tttgc	120
tccctagctg	tcc	cttct	cacc	atacc	ctgc	gttct	gttgc	ccaaagg	180
taaaataaaat	taaa	ataaaat	cacaa	agact	gtgtc	at	attgc	aaat	240
anagtttcat	anag	tttcat	ccg	tttgg	aaaaaacat	tctt	caatt	catctat	300
gcattgtcgt	gcatt	gtcgt	ctat	tttgc	tttgc	tgc	tttgc	tttgc	360
ttgattgtct	ttgatt	gtct	tttcc	at	tttcc	tttgc	tttgc	tttgc	420
tggagaggg	tgg	agag	cc	gg	tttgc	tttgc	tttgc	tttgc	467
<210> 44									
<211> 613									
<212> DNA									
<213> Homo sapien									
<220>									
<221> misc_feature									
<222> (1)...(613)									
<223> n = A,T,C or G									
<400> 44									
tttttttttt	ttttttttt	tttttttttag	ttttaaaata	ttttca	ttttttttttt	ttttaat	ttttaat	ttttaat	60
attccaacag	att	ccaac	actgtt	attttaa	aggc	aggat	gtt	tttca	120
gcagccgggc	gcag	ccgg	cgatt	ttttttt	cagg	acgt	ggcc	tttgc	180
tagctcatac	tag	tcata	actgt	ttttttt	tc	gttgc	tttgc	tttgc	240
cgga	cg	tttgc	atc	ttttttt	tc	gttgc	tttgc	tttgc	300
ctta	ctt	tttgc	tttgc	ttttttt	tc	gttgc	tttgc	tttgc	360
aa	aa	tttgc	tttgc	ttttttt	tc	gttgc	tttgc	tttgc	420
acc	cc	tttgc	tttgc	ttttttt	tc	gttgc	tttgc	tttgc	480
ctt	cc	tttgc	tttgc	ttttttt	tc	gttgc	tttgc	tttgc	540
ttt	ttt	tttgc	tttgc	ttttttt	tc	gttgc	tttgc	tttgc	600
ttt	ttt	tttgc	tttgc	ttttttt	tc	gttgc	tttgc	tttgc	613
<210> 45									
<211> 334									
<212> DNA									
<213> Homo sapien									
<220>									
<221> misc_feature									

```

<222> (1)...(334)
<223> n = A,T,C or G

<400> 45
accagaccaa gtgaatgcga cagggaaatta ttccctgtgt tgataattca tgaagtagaa      60
cagtataatc aaaatcaatt gtatcatcat tagtttcca ctgcctcaca ctatgtgagct      120
gtgccaagta gtagtgtgac acctgtgttgc tcatttccca catcacgtaa gagttccaa      180
ggaaagccaa atcccagatg agtctcagag agggatcaat atgtccatga ttatcaggta      240
tgctgactat ttccaaggggg ttttcagtt gtttcatttgc ttgtaaagc aggtaatcct      300
cttgggtgtnt ttctttttc tcgatgagcc gtgt      334

<210> 46
<211> 429
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(429)
<223> n = A,T,C or G

<400> 46
acaattttnt taaaacaagca gaatagcact aggcagaata aaaaattgca cagacgtatg      60
caatttcca agatagcatt ctttaaattc agtattcagc ttccaaagat tgggtgccca      120
taatagactt aaacatataa ttagtggctaa aaaaaataag tatacggaaa tgtaaaaaaag      180
gaaatgttaag tccactctca atctcataaa aggtgagagt aaggatgcta aagcaaaata      240
aatgttaggtt cttttttctt atttccgtt atcatgcagt ctgtttttt gatatgcctt      300
agggttaccc atttaagtta gaggttgtaa tgcataatggtg ggaatggaaa ttgatcaaat      360
atacaccccttgc tcatttcatt tcaaattgcg gntggaaact tccaaaaaaa gggtaggcat      420
gaagaaaaaa      429

<210> 47
<211> 394
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(394)
<223> n = A,T,C or G

<400> 47
acgcgaantt gtgttatgac ttagatgcctt cagctacaaa angataggac tgacctggtt      60
taaagtgttc tattttgtaa atcattccat tttagtcttt ctgtatgaact tggctatact      120
gaaatctgtt atttttagtga ggctccaaaa tgagcaagc taggcctgtat tagatgttagag      180
tgactattaa aaaacataac ttcttaggag ctataaatca aagttttaaa aagatgtttg      240
gatatatttg agtattccga tcatggaaaact agaaaattgcc ctgcctacta caaggacaga      300
ctgatgggaa attatgcacc tggtcaactt agcttttaag cagacgtgc tgtaaaaaaca      360
aacggcttctt ctgatattta ttgtaaagttt tagt      394

<210> 48
<211> 486
<212> DNA
<213> Homo sapien

<400> 48
acaaaggaaac cgaggggtga ccacctctga gatgtccttgc actttgtcat agcctggggc      60
atattgagca tctctctcac agctgcctt cttatccccca ttcttgatgt agacccctt      120

```

ccgagtcagc	ttttctcct	cctcagacac	aaacagagct	ttgatatcct	gtgcagggag	180
cagctttcc	ttttgttct	ggcaagtgg	agttggagga	gcctcaaag	ctcgagttgt	240
tcctcggt	caggggagac	aatgggcct	gatgtctgg	ccatattca	gcttattctt	300
gagttgatc	agggcaacgt	catagtcata	aaattcagga	attcctgtt	ctttttccc	360
attaatgtt	tagttgggt	gaaataggac	tacttctatc	tccaggcccc	gcttctcccc	420
tcccttgatt	gagtgttct	tgtcatccac	agtgaaacaa	tgtgctgtg	teagcacaaa	480
gtacct						486
<210> 49						
<211> 487						
<212> DNA						
<213> Homo sapien						
<400> 49						
acgggctgac	agagaagatt	cccgagagta	aatcatctt	ccaatccaga	ggaacaagca	60
tgtctctcg	ccaagatcca	tctaaactgg	agtgatgtt	gcagaccagg	cttagagttc	120
ttctttctt	cttaagccct	ttgctctgg	ggaagttctc	cagttcagc	tcaactcaca	180
gcttctccaa	gcatcaccc	gggagttcc	tgagggttt	ctcataaatg	agggctgcac	240
atggctgtt	ctgcttcgaa	gtattcaata	ccgctcagta	ttttaaatga	agtgattcta	300
agatttgggt	ttggatcaat	agggaaagcat	atgcagccaa	ccaagatgca	aatgtttga	360
aatgatatga	ccaaaatttt	aagttagaaaa	gtcaccggaaa	cacttctgt	ttcaacttaag	420
tgtctggccc	gcaatactgt	aggaacaagc	atgatcttgt	tactgtgata	ttttaaatat	480
ccacagt						487
<210> 50						
<211> 460						
<212> DNA						
<213> Homo sapien						
<220>						
<221> misc_feature						
<222> (1)...(460)						
<223> n = A,T,C or G						
<400> 50						
acatattttg	gttgaagaca	ccagactgaa	gtaaaacagct	gtgcacccaa	tttattatag	60
ttttgttaat	aacaatatgt	aatcaaactt	ctaggtgact	tgagagtgg	acccctata	120
tcattattta	gcaccgttt	tgacagtaac	catttcagtg	tattgtttat	tataccactt	180
atatacaactt	attttcacc	aggttaaaat	tttaatttct	acaaaataac	attctgaatc	240
aagcacactg	tatgttcagt	aggttgaact	atgaacactg	tcatcaatgt	tcagttcaaa	300
agcctgaaag	tttagatcta	gaagctggta	aaaatgacaa	tatcaatcac	attagggaa	360
ccattgttgt	cttcacttaa	tccatttagc	actattgaaa	ataagcacac	caagntatat	420
gactaatata	acttggaaaat	ttttataact	gagggggtn			460
<210> 51						
<211> 529						
<212> DNA						
<213> Homo sapien						
<400> 51						
acacttggaaa	ccaaatttct	aaaactgtt	tttcttaaaa	aatagttgtt	gtaacattaa	60
accataacct	aatcagtgtt	ttcaactatgc	ttccacacta	gccagtcctc	tcaacttct	120
tctggttca	agtctcaagg	cctgacagac	agaagggttt	ggagattttt	tttcttaca	180
attcagtctt	cagcaacttg	agagtttct	tcatgttgc	aagcaacaga	gctgtatctg	240
cagttcgta	agcatagaga	cggtttgaat	atcttcagtg	gatatcggt	ctaaactgtca	300
gagatgggtc	acaaaacata	atcctgggg	catactggcc	atcaggagaa	aggtgtttgt	360
cagttgtttc	ataaaaccaga	ttgaggagga	caaactgctc	tgccaaatttc	tggatttctt	420
tattttcagc	aaacactttc	ttttaagctt	gactgtgtgg	gcaactcatcc	aagtgtatgaa	480

taaatcatca	agggtttgtt	gcttgtcttg	gatttatata	gagcttctt	529
<210> 52					
<211> 379					
<212> DNA					
<213> Homo sapien					
<400> 52					
actttgccaa	gcagtaaagg	atccaggaga	tagcactgga	tgtgggtcgtca	tgtcctgcaa
acatgaacgt	tttcacttca	gcctggagat	ctgcttcaga	gaaatctttg	gtgtttcgc
ttttggact	caaaagtatg	tccagaaaat	cccagcgcct	tttctgagta	gtatcttgg
tttagtttatc	cttaagagac	tccttccgg	cctggattac	tttctctgtg	aactgatgaa
gttcttggtt	aaattttagaa	aagatttggc	cttgagagct	gaatttggaaa	accaggtcgt
tgtgatgtag	aaaatttggc	atgcgtgg	tggagatttt	gctaagggtt	aacactgctt
tcaggatga	gtccagggt				379
<210> 53					60
<211> 380					120
<212> DNA					180
<213> Homo sapien					240
<220>					300
<221> misc_feature					360
<222> (1)...(380)					379
<223> n = A,T,C or G					
<400> 53					
acttttatct	taaaagggtg	gtagtttcc	ctaaaatact	tattatgtaa	gggtcattag
acaaatgtct	tgaagtagac	atgaaattt	tgaatggttc	tttacattt	ctctcccccc
tttttggcat	cctggcttgc	ctccagttt	aggccttta	gtttgcttct	gtaagcaacg
ggaacacactg	ctgagggggc	tcttccctc	atgtataactt	caagtaagat	caagaatctt
ttgtgaaatt	atagaaattt	actatgtaaa	tgcttgatgg	aatnntttcc	tgctagtgt
gcttctgaaa	ggcgctttct	ccatttattt	aaaactaccc	atgcaattaa	aaggtaacctt
cccgccgacca	cnctaanggc				380
<210> 54					
<211> 245					
<212> DNA					
<213> Homo sapien					
<400> 54					
gcgccggcgct	tcacttcttc	aacttccgg	ccggctcgcc	cagcgcgctg	cgagtgtgg
ccgaggtgca	ggagggccgc	gctggattta	atccaaaaga	gggatgtaaa	gttcacgtgg
tcttcagcac	agagcgctac	aacccagagt	cttacttca	ggaaggtgag	ggacgtttgg
ggaaatgttc	tgctcgagtg	ttttcaaga	atcagaaacc	cagaccaacc	atcaatgtaa
cttgt					245
<210> 55					60
<211> 556					120
<212> DNA					180
<213> Homo sapien					240
<400> 55					245
acagaagatg	aataataatg	aaaaactgtg	atttttgac	tatcacatac	attgtgttaa
aaaacaggtt	aatataatga	ctattactgt	taagaaaagac	aaggaggaaa	actgtttcaa
tgttcaggtt	taaataactaa	gcacaaaaat	ataacaaaatt	ctgtgtctac	aataattttt
gaagtgtata	caagtgcatt	gcaaatgagc	tctttaaaat	ttaaagtcca	tttccctttt
agccaagcat	atgtctacat	ttatgattt	tttctcttat	ttaaagtct	tttctgggtt
					300

agtttttaa aaagtttcat catggctgtc atcttggaat ctggctcca gctcaaagct	360
gagacttcac gcatacatat ttcctttct gttgcacatc tcacctagtt tctccaagta	420
ttcagagtt aatagcacaa cttctttat atgttcacatt ttgtccacat gtatgtgcag	480
tgctgctgtc tcaatgttgc ttctcacaca ccctttccct tcttcaaca gcagtcacca	540
aacgttcaca acacaa	556
<210> 56	
<211> 166	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(166)	
<223> n = A,T,C or G	
<400> 56	
atggcccttg attacatcat tatgaactac tcaggnnaac atccaaata ccgacctngg	60
gaaagacttg gtccgagatg tgttcatcca tacaggctac ctttccaga gcncaggnc	120
caagagctgc ntatcacct acctggccca ggtggacccc anaggg	166
<210> 57	
<211> 475	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(475)	
<223> n = A,T,C or G	
<400> 57	
acatccncat gttcctccaa atgacgttg ggtcctgct tgccaacatt ctttattgcc	60
agctgttcag gtgtcatctt atctttcttct tctacagcct tattgttaatt cttggctaat	120
tccaaacatct ctttaccac tgattcattt cggttacaat gttcactgtt gtcctgaagt	180
gtcaaacctt ccatccaaact cttcttatgc aaatttagca acatcttctt ttccagttca	240
ttttccgat agttaatagt aatggagtaa taatgtctgt ttagtccatg aattaatgcc	300
tggatagatg gttgtttaa gtgaccacaga ttcaagttt tttgtctgg ttcatgttct	360
aagaccatca tattagcatt gatcaatctg aaggcatcaa taacaaccc ttcccttaca	420
ctctgaatgg gatccacaaac cactggccaca gntctctccg ataaggcttc aaagc	475
<210> 58	
<211> 520	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(520)	
<223> n = A,T,C or G	
<400> 58	
actgttnatg tgctacttgc atttggccctt cttcctgtgc actaaagaccc ccactcactt	60
cccttagtgtt cagcagtggaa tgacctctag tcaagaccc ttgcacttagga tagttaatgt	120
gaaccatggc aactgatcac aacaatgtt ttcagatcg atccattttta tcctccttgc	180
tttacagcaa gggatattaa ttacatgtt taccttccctt tggactatg aatgtgcaaa	240
attccaatgt tcatggtctc tccctttaaa cctatattctt accccttttta cattatagaa	300
aggaatgctg gaaacccaga gtccttctt tggacttctt aatgtgttatt tctaattatc	360

catgactctt aatgtgcata tttcaattg cctaatngat ttcaattgtc taagacattt	420
caaatgtcta attggggaga actgagtctt ttatatcaag ctaatatcta gctttatat	480
caagctaata tcttgacttc tcagcatcat agaagggggt	520
<210> 59	
<211> 214	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(214)	
<223> n = A,T,C or G	
<400> 59	
ctggcaggaa atgcatcaaa agacttaaag gtanagcgta ttacccctcg tcacttgcaa	60
cttgcatttc gtggagatga agaattggat tctctcatca aggctacaat tgctgggtgn	120
ggtgtcattc cacacatcca caaatctctg atnggaana aaggacaaca naagactgnc	180
taanggatgc ctgnatncct tggaaatctca tgac	214
<210> 60	
<211> 360	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(360)	
<223> n = A,T,C or G	
<400> 60	
gcataacaaca tggcagcagg gcctcggaa gangggtagg aggaccgagc agcattctct	60
gtagaggaag acaggaaagg agacccctttt ggcacacatt tatggaggggt tgccctgaa	120
gagaaggcga ggtgggagag gttccctgtt acttaagaga aggcaccagt ggcaaagagc	180
acaatgaaga ggatgtatgt aaaaacaatc acgcagataa ggacaatcat cttcacgttc	240
ttccaccaga attttcgagc caccttctgc gatgtcgctc tgaagtgctc agatgtggct	300
tccagatcct ctgtcttgtt gcggagatgt tccaagttt cccccccggc caggatccgc	360
<210> 61	
<211> 391	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(391)	
<223> n = A,T,C or G	
<400> 61	
tntgggatcg tactcgatta aacagagcca cctttgttcc tgaggcaatg cataantcan	60
cattttcaa tgactgcttc ttttggaaag gnttggagat gacttttatac cgcttgctga	120
ggaacacacc aatgnatca ctgttgcattt agaacatctt tacagacaac atgaantgct	180
ttcgcttgctc tgagtcagat atatacaatg ttttggctgt gcaataggtc ttcccttcca	240
agtttagctg ctgcatttct tggncactat ttccatccc aataaatgca cacggttgag	300
actcttgntc agaacaacca tcncgttcca tttgttctt tttntcttc catccactgc	360
ccataagata tacacannga ggtggcaaa a	391
<210> 62	

```

<211> 324
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(324)
<223> n = A,T,C or G

<400> 62
acaattttat tttaacagat ttcaagagtc catttttaa aaaatgagca ataaagaacc      60
tctatcagtg agacttctca ttttatagca aatacatttt tgcagcttaa attttcttga      120
attcatatac gcttctgtca ttaaacaaaa cttccagaga aaactggct ctatataattt      180
aagtaacaaa ttgacaaaa tacatattta tacatataata ganctcta ataaatatta      240
aatttgaaaa aatcaaatgt gaagcagaaa ctgctataca agtataattgt ntaatatcta      300
tntnatacat taaagnnntc cggg      324

<210> 63
<211> 360
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(360)
<223> n = A,T,C or G

<400> 63
acagannccct tgaatatgtt gtgggtccct cattatggcc cttcattccc ttctgtgtta      60
atagtaaaggc atgttgccta ataactacaa ccctgaccaa atttgggcct ggatctcatg      120
ggtcacgtgg agtttaat acgatttttta atttacttgg gtaattgagc tgaatcttta      180
gtttcagat tactttttta aacagatagg ctcttagaac aaattatataa aaacataata      240
cccatttggc ggggaatctg gattaactac coactgttcc caccgggggg aactttgaa      300
aaattttggc catatagaat gcatggaaaaa tcaggtatga tcttatgagg actttatagt      360

<210> 64
<211> 491
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(491)
<223> n = A,T,C or G

<400> 64
nctgactgtg atgtccactt gttccctgat ttttacacat catgtcaaag ataacagctg      60
ttcccaccca ccagttccctc taagcacata ctctgctttt ctgtcaacat cccatttgg      120
ggaaaggaaaa agtcatattttt attcccgac cccagttttt taacttggcc tcccaacttgc      180
cccccttcc tctgggtgtt agaaggaaaa ttggaaaaaa atttatataa tatttcttcc      240
ttaatgggtgg ggggctactg gagaggagag acagcaagtc caccctaaact tgttacacag      300
cacataccac aggttctggc attctcatct tcgaacctag agaaataggt gctataaaca      360
gggaattaag caaaatgctg gatgtatag atcttttaat tgncttaatt tttttcttat      420
tattaaacta caggctgttag atntcttagg tctcacagaa ctnttatcat tttaaactga      480
cttgatattt t      491

<210> 65
<211> 484

```

```

<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(484)
<223> n = A,T,C or G

<400> 65
accagcacac cggcgccgtc ctggactgct ctttctacga tccaaacgcat gcctggagtg 60
gaggactaga tcatcaattt gaaatgcattt atttgcacat tgatcaagaa aatcttggtt 120
ggacccatgtc tgccctatc agatgtgtt aatactgtcc agaagtgaat gtgatggta 180
ctggaaaggttt ggatcagaca gctaaactgtt gggatcccag aactccttgc aatgctggga 240
ccttctctca gcctgaaaag gtatataccct tctcgtgtc tggagacccg ctgattgtgg 300
gaacagcagg cccgagagng ttggtgtggg acttacggaa catgggttac gtgcagcagc 360
gcagggagtc cagcctgaaa taccagactc gtcgtatc aacgtttcca aacaagcagg 420
gttatgtatt aagctctattt gaaggccgag tggcagttga gtatggac ccaagccctg 480
aggt 484

<210> 66
<211> 355
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(355)
<223> n = A,T,C or G

<400> 66
ngaagaagaat atgggtggag gtgaaggtaa tcacagagct gtcgtattctc aaaacagtgg 60
tgaaggaaat acagggtctg cagaatcttc ttttctcag gaggtttcta gagaacaaca 120
gccatcatca gcacatgtaaa gacaggcccc tcgagcacctt cagtcaccga gacgcccacc 180
acatccactt cccccaagac tgaccattca tgcccccacctt caggagttgg gaccaccagt 240
tcagagaattt cagatgaccc gaaggcagtc tgcgttgcgtt ggccttcagt tgactccagg 300
aataggtggc acgcaacacgc attttttga tgatgaagac agaacagttc caagt 355

<210> 67
<211> 417
<212> DNA
<213> Homo sapien

<400> 67
acgacaccccc tcaagagggtt gcccgaagctt tcctgtcttc cctgacagag accatagaag 60
gagtcgtatgc tgaggatggg cacagccccag gggaaacaaca gaagcggaaat atcgttctgg 120
acccttcagg ctccatgttac atctacctgg tgcttagatgg atcagacacgc attggggcca 180
gcaacttcac agggagccaaa aagtgtcttag tcaacttaat tgagaagggtt gcaaggatgt 240
gtgtgaagtc aagatgttgc tctgtacat atgcccacata ccccaaaaattt tgggtcaag 300
tgtctgaagc agacacgcgtt aatgcacactt gggtcacgaa gcagctcaat gaaatcaattt 360
atagaagacca caagttgaag tcaggacta acaccaagaa ggcctccag gcagtgtt 417

<210> 68
<211> 223
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature

```

<222> (1)...(223)
 <223> n = A,T,C or G

<400> 68
 cacttgcaag ctgtttaca gagacctgnt aaacaaagaaa cagacagatt ctataaaatc
 agttatataca acatataaaag gagtgtgatt ttcatgttgc ttttttaagt aaatatgacc 60
 aactgacta aataagaagg caaaacaaaa aattatgctt cttgacaag gcctttggag 120
 taaacaaaaat gcttaaggc tcctggtaa tgggttgca agg 180
 223

<210> 69
 <211> 396
 <212> DNA
 <213> Homo sapien

<400> 69
 acctttttc tctccaaagg aacagtttct aagtttctt gggggaaaaaaa aaaacttaca
 tcaattaa accatatgtt aaactgcata ttatgtgt tacacaaaaa aattgcctca 60
 gctgtctac acaagttca aagtattaa tgcttgat aatattactc aacattaaat 120
 tattttaaat tattattaa aaaaaaaaaact ttcttaaggaa aaataaacaa attagacccg 180
 tgattatcaa aggattat aagaatctt accaaaaatt tcaaccctac aacctaaaac 240
 cgcaaatttc tattttaaat catcagaaaa taactcttgc ttcttactt atgacccaaa 300
 gtttttattt cactattcaa tatctgaaaaa gtatca 360
 396

<210> 70
 <211> 402
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(402)
 <223> n = A,T,C or G

<400> 70
 acccannccc acccaggcaa acagctccga catgtttngt aagttagaca agccagtgc
 agttttttt tttttttctt tttttttttt ttgtctttt gcttacccctt ttgcttaatg 60
 gaattgttat ggctaaggcac atagaaggcc aaaaaaggag ttttcaaac ccagcaatc 120
 aagtgttgg atttgcactt gccaaaaagaa aactgcactt cccctctaa gtaaaacgaa 180
 atgagtttctt taggttaatg tattcatcag cccagataaa aaaaaaaaaacca gttatgtgag 240
 cgttagtcac tgctcatttc caggaanatc aaacaaaata ccagcccagc cagactcaca 300
 tgtgggnata tatataaaaa gcaagagagc cacacccaca ag 360
 402

<210> 71
 <211> 385
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(385)
 <223> n = A,T,C or G

<400> 71
 accagtagag agtggccctt gcaggccact tataaacagg aagctctctc ctgagctcac
 tgatcaacctt gcccggca cagacagaac ctaccagaaa agaacaagta caaaaacta 60
 tcattatctg tttctcaag acagtccaa atgtcttgc tgcgtcgcca caaaactcagt 120
 gattggccca agtcattccc ggggtccata aacagtaact ggtgtgcanc attagaacaa 180
 ggggacacggc ctttgcattctt cttctgagca acatgaactg ggatttctgc cnccccggat 240
 300

ctcggtgcc	acctccgaag	aagtctgac	cagccaccc	cacagtaaaa	gattcctccc	360
gtgagatga	tttggatgc	gncc				385
<210> 72						
<211> 538						
<212> DNA						
<213> Homo sapien						
<220>						
<221> misc_feature						
<222> (1)...(538)						
<223> n = A,T,C or G						
<400> 72						
caattaatta	acagaggat	aattgtctca	cttcagaag	tgatcattta	tttttattta	60
gcacaggta	taagaaaaat	ataagaaaa	ataatcaatt	tcatatataa	aaggattatt	120
tctccaccc	taattattgg	cctatcatt	gttagtgtt	tttggtcata	ttattgaact	180
aatgtattat	tccattcaaa	gtcttctag	attaaaaat	gtatgcaaaa	gcttaggatt	240
atatcatgtg	taactattat	agataacatc	ctaaaccc	agtttagata	tataattgac	300
tgggtgtaa	ctctttgt	atctgnntt	acagatttct	taaattatgt	tagcataatc	360
aaggaagatt	taccttgaag	cacttccaa	attgataactt	tcaaaacttat	ttaaaaggcag	420
tagaacctt	tctatgaact	aagtccatg	caaaactcca	acctgttaagt	atacataaaa	480
tggacttact	tattcctc	accttctcca	ggccttagaa	tattcttctc	tggagccc	538
<210> 73						
<211> 405						
<212> DNA						
<213> Homo sapien						
<220>						
<221> misc_feature						
<222> (1)...(405)						
<223> n = A,T,C or G						
<400> 73						
actttatnna	tgaaattttc	ttctacttgt	atccattnc	cggggcttat	ggacccattc	60
atactctcca	tattnaaat	caaaggttcc	tttctgaaga	gacctaatt	ttaaggtaaa	120
acgtggtcca	agttccctaa	ttcccaactt	cttttcaactc	ctgaatatgt	atctgtgaaa	180
tctgaagaat	atgtaatccc	gttgattgtg	gaatgtggca	acctgccttc	cgataaattg	240
aggattatga	ggaaagagag	atgcaaacat	acgtccaaatt	gaatgaccca	gccgtgttgt	300
aaaattatcc	agaattattt	caggtatgtg	ttctgtgggg	tccttgctc	ttctcttaat	360
ttcttacga	agacgaacac	tgctcattt	aaaatgagca	gttgg		405
<210> 74						
<211> 498						
<212> DNA						
<213> Homo sapien						
<220>						
<221> misc_feature						
<222> (1)...(498)						
<223> n = A,T,C or G						
<400> 74						
tgagccctgc	acctgtttcc	tgcacccccc	gccnactgg	tctatggcca	caaggagttt	60
tacccagtaa	aggagttga	ggtgtattat	aagctgatgg	aaaaataccc	atgtgctgtt	120
cccttgggg	ttggaccctt	tacgatgtt	ttcagtgcc	atgaccaga	ctatgccaag	180
attctcctga	aaagacaaga	tcccaaagt	gctgttagcc	acaaaatcct	tgaatcctgg	240

gttggtcgag gacttgtgac cctggatgg tctaaatgga aaaagcaccc ccagattgt	300
aaacctggct tcaacatcg cattctgaaa atattcatca ccatgatgtc tgagagtgtt	360
cgatgtgc tgaacaatg ggaggaacac attgcccac actcacgtct ggagctctt	420
caacatgtct ccctgatgac cctggacac atcatgaagt gtgccttcag ccaccaggc	480
agcatccagt tggacagt	498
<210> 75	
<211> 458	
<212> DNA	
<213> Homo sapien	
<400> 75	
agccttgcac atgatactca gattcctcac ctttgcttag gagtaaaaca atatactta	60
cagggtgata ataatctcca tagttatgg aagtggctt aaaaaggca gattgacttt	120
tatgacatttgcataaaatct acaaatcgc cctcgagttt tcaatgata actgacaac	180
taaatttgcattt cccttagaaag gaagatgaaa ggagtggagt gtggttggc agaacaactg	240
catttcacag ctttccatgt taaattggag cactgaacgt tcagatgcat accaaattat	300
gcatgggtcc taatcacaca tataaggctg gctaccagct ttgacacagc actgttcatc	360
tggccaaaca actgtggta aaaacacatg taaaatgctt ttaacagct gatactgtat	420
aagacaaagc caagatgca aattaggctt tgattggc	458
<210> 76	
<211> 340	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(340)	
<223> n = A,T,C or G	
<400> 76	
accttataacc aaaanaatgc ttattccaaa atatttttg tagcttagtag ttcttcctt	60
ggaggtaag aaaatacacc caaacttta attaccagga ttcaagatata ttaagagaac	120
aatttttagtt aagaatcaaa tatactgaga ttcaaaagggg ggaaaaaaaaag gaaatattat	180
agaagacaaa ggtcaaactg gcattccaga tctggagcaa ttttgtaaag caggaaaaca	240
actatgacaa tctgnagctt cttagatcat tatagtgaat gtnccattt actataaggg	300
tttttataat ggtgtttcct aaataaagga acataaatgt	340
<210> 77	
<211> 405	
<212> DNA	
<213> Homo sapien	
<400> 77	
actccatttg tggaaactcggt gtcggagtcgt ggtaaacagc cgaatgtctt cctccctac	60
agtttcctct ctttgcatttgc gaggcgtat gtctgtt aaggcattaa ttttatctat	120
caggaagaac attttttcat tttcgcttc cggatgtcg acaccataact ttttagtctc	180
ctctgttatt ctctggtagt ttcctgtat ttgattttct aacagggggca gagattaca	240
gataatgtgtg atgagctcg tggtaagttt ttctgcccagg cagggaaaccc tgcccttcc	300
ttccctccagc agatccctga aatatgggtt gttctcaaag aagatcttct ctctctgcag	360
ggcttcggac aggctcagct ggtctggat ctctgttgg ccccg	405
<210> 78	
<211> 410	
<212> DNA	
<213> Homo sapien	

```

<220>
<221> misc_feature
<222> (1)...(410)
<223> n = A,T,C or G

<400> 78
acagcagntn tagatggctg caacaacatt cctccttaccc cagcccagaa aatatttctg      60
ccccacccca ggatccgggg cccaaaataaa gagcaaggag gcccccttca ctgagggtgct    120
gggttagggct cagtgcacca ttactgtgct ttgagaaaaga ggaaggggat ttgttggca    180
ctttaaaaat agaggagtaa gcaggactgg agaggccaga gaagatacca aaattggcag    240
ggagagacca tttggcgeea gtccctttagg agatgggagg agggagatag gtatgagggt    300
aggcctaag aagagttagga ggggtccact ccaagtggca gggtgctgaa atgggctagg    360
accaacagga cactgactct aggttatga cctgtccata cccgttccac      410

<210> 79
<211> 512
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(512)
<223> n = A,T,C or G

<400> 79
acagtgaaaaa acaaactaat ataaagcatt ccagnngata aaaaccttcc caggctttag      60
gtttgttttc caaggaaatt atgttcaat gtaaagtttgc aaatactcca gacatacatt    120
ccatgttaggt tttgggtgcc aatgttaaaa ttcaaattt tgcattgcaag gcttagcaaa    180
gaaacactgg cagaattcca gcatttgc aattctaagt tttgggtgaat attgttaata    240
ttacaatttgg tattagaaag ccatgatgaa tccagaatta agagaaaaacc catttcataa    300
atattttgtt tgattaaaaaa ataccaggct taccatgttc taaataacac aagaaaaatat    360
ctttaaaaaaa aaaaggactg caatttaaca gtaatctgta tatcttttagc tgccattaaa    420
aaaagaaaaaa agaacaacca aaaacaatga aaatgttaca actggtataa agtnaccnna    480
tgatgctccc cttacgagaa aacaaaactg tc      512

<210> 80
<211> 174
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(174)
<223> n = A,T,C or G

<400> 80
tgattccca gacctcaaat gggctaacac gtttcttcc tncagcagnc ttccgtccg      60
tgaagtncc ttccagattt gtacatggaa ctgaaaacaa agggagcctc agctggattt    120
aaatctggag catgccacaa agncttgcac tnggcatttt cnagaagaac ccat    174

<210> 81
<211> 274
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(274)

```

```

<223> n = A,T,C or G

<400> 81
ttgcaacaag cacattaaat taaggcctgc tngaatttct tcctccccaa tcaggtaaac      60
tttcttgc aataaaagttt gaggaggtgg catttggaaa ttctttaaa aaagaagtct      120
tcatctattc acnagaaaaac tcaaaaataa ttttcattat caacacacaa actaactcaa      180
tctctgttt aagtttctat tggcaattt ttctgattna tacgagaatt attntcagnt      240
ntagaaaatc ctggtctttg gtcattacaa gntg      274

<210> 82
<211> 101
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(101)
<223> n = A,T,C or G

<400> 82
atggagaaga tcgaacctga gcctnntgag aattgcctgc tacngcctgg cagccctgcc      60
cgagtggccc agcnnccattt cacnagntgg gcatgatttg n      101

<210> 83
<211> 182
<212> DNA
<213> Homo sapien

<400> 83
tattatgggg aaagataact gagaataaaag ctatcatgca gatatttgc gagataaaag      60
taatgcagat actgagtgg a gttttgatca aactatgctt gaaagccact ctaccactag      120
ttcacacaaac caataatttc ctttcgactt ggaagtcagc ttgagtttt tcaggtgttt      180
tt      182

<210> 84
<211> 229
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(229)
<223> n = A,T,C or G

<400> 84
actgtttgtt gctgcactac aacagattct taccgtctcc acaaaggtaa gagattgtaa      60
atggcataa ctgactttttt ttttattccc ttgactcaag acagctaact tcattttcag      120
aactgtttta aacctttgtt tgctggttta taaaataatg tggtaatcc ttgtttgttt      180
cctgatacca nactgtttcc cggngttgg tagaatatat tnnngttcng      229

<210> 85
<211> 500
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(500)

```

<223> n = A,T,C or G

<400> 85

ggggagtang tgatttatta aagcaagacg ttgaaacctt tacnttctgc agtgaagatc	60
aggggtgtcat taaaagacag tggaaaccag gatgaaagtt tttacatgtc acacactaca	120
tttcttcaat atttcacca ggactccgc aatgaggcgtt cgtttctgaa gggacatctg	180
atccgagcat ctcttcaact ctaacttggc tgcaacagct tccagagggg catcaaattt	240
ggcaagactt aacttgaaca gaggttcaact aatgaagaag aagtctaaca gctcagaaac	300
aagagctggg cagaactcggtt cattggcctg gtacgcgcg agggccagcg tgaccagcag	360
gagacacacc gacagcttca tggtggcttg tttgctgtg agctcagtt tcacaaacaa	420
tgagtgtttt ggactccacc ccaggagcgtt gttggagctgc agagcccagg gctatttgc	480
cctgccccggg cggncgctcg	500

<210> 86

<211> 323

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(323)

<223> n = A,T,C or G

<400> 86

ccgcccagtgt gctggattc gcccattggccg cccgggcagg tactcagaag tcatttgtta	60
tttacaattt ggtttgtgtt ggtatggatn tangggat gggccatgtc ttttgcattt	120
aagatgcaat antcattgtc ctctcccaact gtctccctttt tcctcaccctt atggcagctn	180
tcatgaccctt ttcccaaagg gtccaccggag tcctgaactc agtttcatca ccaacattcc	240
tcgccttcagg ttgaattcaa cactgncaan ggagnagang caaagacttg ggtcagggag	300
aggggggaa acacanaaca aac	323

<210> 87

<211> 230

<212> DNA

<213> Homo sapien

<400> 87

gcagcatttga gccacccctt tggcaggcga tacggcagct ctgtgcctt ggcacatgt	60
tggagtggag gagatgctgc ccctgtgtt ggaacatctt ggggtgaccc ccgacccagc	120
ctcgctgggc tggccctgtt ccctatctct cactctggac ccagggtctga catcctaata	180
aaataactgt tggatttagac aaaaaaaaaa aaaaaaaaaa aaaaaaaaaagg	230

<210> 88

<211> 249

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(249)

<223> n = A,T,C or G

<400> 88

atgtgaccag gtcttaggtct ggagtttcag ntggacact gagccaagca gacaagcaaa	60
gcaaggccagg acacaccatc ctgccccagg cccagcttctt ctccctgcctt ccaacgccc	120
ggggagcaat ctcagccccca aactctgcctt gatgcctttt atcttggcc tctttgtctgg	180
aggtgtgacc accactccnt ggtcttggc ccggccccat ggatcctgtc ctctggaggg	240
ggtntagat	249

```

<210> 89
<211> 203
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(203)
<223> n = A,T,C or G

<400> 89
tgtttacact gtcaggatg acaaggaaag tggcntatc tntgataccca tcatcccagc      60
tgttccctt cccactgacc tgcgattcac caacattggt ccagacacca tgctgtcac      120
ctgggctcca ccccccataa ttgatttaac taacttcctg gtgcgnact caccctgtcaa      180
aaatgangaa gatgttgcag agt      203

<210> 90
<211> 455
<212> DNA
<213> Homo sapien

<400> 90
ctctaagggg gctggcaaca tggctcagca ggcttgcccc agagccatgg caaagaatgg      60
acttgtaatt tgcatttcgg tgatcacctt actcctggac cagaccacca gccacacatc      120
cagattaaaa gccaggaagc acagcaaacg tcgagtgaga gacaaggatg gagatctgaa      180
gactcaaatt gaaaagctt ggacagaagt caatgcctt aaggaaattc aagccctgca      240
gacagtctgt ctccgaggca ctaaagttca caagaaatgc tacttgcctt cagaaggttt      300
gaagcatttc catgaggcga atgaagactg catttccaaa ggaggaatcc tggtatccc      360
caggaactcc gacgaaatca acgcctcca agactatggt aaaaggagcc tgccagggtt      420
caatgacttt tggctggca tcaatgacat ggtca      455

<210> 91
<211> 488
<212> DNA
<213> Homo sapien

<400> 91
actttgcctt ctcatatgca tggtagtcaact ttataagtca ttgtatgtta ttatattccg      60
taggttagatg tggtaacccctt tcacccattt catggctgaa gtcacccctt gttacagta      120
gcgtacgtg ggcgtgtgca tggccttgc gctgtgacc accacccaa caaaccatcc      180
agtgacaaac catccagtgg aggttgcg ggcacccagcc agcgtacgtg ggtcggaaa      240
ggccacctgt cccactcata cgtacgcta ctataaagag aagacgaaat agtacataa      300
tatattctat ttttatactt ttccatattt tggtagtgc accgttgcg tggatgtttt      360
tctacccaaac ggcctgcag ccagtcacg tccaggttca acccacagct acttggttt      420
tggttcttcatattctaa aaccatttca ttccaaagca ctttcagttcc aataggttta      480
ggaaatag      488

<210> 92
<211> 420
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(420)
<223> n = A,T,C or G

```

<400> 92
 tctccggcag gctctgcccc ggtcgtagcn agnnaaccta taatcctgac ctttttgta 60
 gacaacccctt gtcgtgaggt taactccatc cattgttagtgcctgtatataatgggacg 120
 attgcattt ttcctgggt gagcttcca gaggtctgaa attttctccc caccttagt 180
 ctgagataact ttatcatgat cgancactc cgccactcc acgtnttggaa cccactca 240
 ggacaaaagaa acattgaaat attcgcatg cttctgtctgg aacaatttga atacccggc 300
 agcagcagag cctcgatgnc caggatattc aatatggtct tccactgaag atgatggatt 360
 tcctttcaca gntagaaaac ttncnagggn gtctaaatcc aaggtgcagg aangngngc 420

<210> 93
 <211> 241
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(241)
 <223> n = A,T,C or G

<400> 93
 accacgaatt ncaacatcca gatccaccac tattctaatggattgttaacttgngactgt 60
 gccccgctcc tgaaagccga ccaccatgca accaacgggg tggtgccacct catcgataag 120
 gtcatctcca ccatcaccaa caacatccag cagatcattt agatcganga caccttgag 180
 accttcggg ctgctgnggc tgcattcaggctcaacacga tgcttgaagg naacggncag 240
 t 241

<210> 94
 <211> 395
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(395)
 <223> n = A,T,C or G

<400> 94
 actctattttt aattctgcct ttttatactt aattctaaat ttttccctc taatttacaa 60
 caaattttgt gatTTTTATA agaatctatg cctcccaat tctcagattt ttctctttc 120
 tccttttattt ctttgcttaa attcagttttaa agctttcttgg tatttttagg ctttcatgcac 180
 attcttattt ctaaacacca gcaggcttc agagacctaa aatccagttt aggaataact 240
 gtgttagttc ttgaaaagac attaaagaca ttttccctg aaacatacag aacatgtcat 300
 gccaaatctc ttgtttacat aataaaactgg taataccggtaaattgcaca tacagattt 360
 atctccaaga tagaataact taaatattaa aacgt 395

<210> 95
 <211> 304
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(304)
 <223> n = A,T,C or G

<400> 95
 cgaggatcag tgatngctcc ccctggccaa tacaatacaa gaacngnggg ttttgcataaa 60
 ttggaaacaag gaaacagaac cacagaaata aatacattgg ttaacatcg attagttcg 120

gttactttt	tgtaaaagtt	aaagtacgag	gggacttctg	tattatgcta	actcaagtt	180
actggaatct	cctgtttct	ttttttttt	taaatnggtt	ttaattttt	ttaattggat	240
ctatcttctt	ccttaacatt	tcagttggag	tatgttagcat	ttagcaccac	tggctnaaac	300
ctgt						304
<210> 96						
<211> 506						
<212> DNA						
<213> Homo sapien						
<400> 96						
acaactgtcag	cagggactgt	aaacacagac	agggtcaaag	tgtttctct	gaacacattt	60
agtggaaatc	actgtttaga	acacacacac	ttacttttc	tggtctctac	cactgctgat	120
attttctcta	ggaaatatac	ttttacaagt	aacaaaaata	aaaactctta	taaatttcta	180
tttttatctg	agttacagaa	atgttactg	aggaagagatta	ctcagtaatt	tgtttaaaaaa	240
gtataaaaat	tcaacaaaaca	tttgctgaat	agctactata	tgtcaagtgc	tgtcaagggt	300
attacactct	gttaattgaat	attattcctc	aaaaaattgc	acatagtaga	acgctatctg	360
ggaagctatt	ttttcagtt	ttgatattc	tagtttatct	acttccaaac	taatttttat	420
ttttgctgag	actaatctta	atcattttct	ctaataatggc	aaccattata	accttaattt	480
attattaacc	ataccctaag	aagtac				506
<210> 97						
<211> 241						
<212> DNA						
<213> Homo sapien						
<220>						
<221> misc_feature						
<222> (1)...(241)						
<223> n = A,T,C or G						
<400> 97						
attttctttt	taattacttt	agagagctag	ggatgcaaatt	gttttcagtt	agaaagcctt	60
tatttacttt	tggaaatgtt	acaagaaatg	catctgtctt	agaaaactgga	gattatttga	120
tgttaggtaa	aacatgtaat	tgtntctctg	gcaaatttgt	atcannattt	ngaaaatgag	180
atattangaa	aaacccaattt	ttcttaaattc	tagnnncatct	ttctttanaa	gaacatttana	240
t						241
<210> 98						
<211> 79						
<212> DNA						
<213> Homo sapien						
<220>						
<221> misc_feature						
<222> (1)...(79)						
<223> n = A,T,C or G						
<400> 98						
ggcaaacana	cttatgctgn	ancngggttt	tancaagggtt	ttcaaaagnaa	aaancccatt	60
ngactttatg	gaaaatattt					79
<210> 99						
<211> 316						
<212> DNA						
<213> Homo sapien						
<220>						

```

<221> misc_feature
<222> (1)...(316)
<223> n = A,T,C or G

<400> 99
ccacatatgt aaaacccaga aagaccngnt tngcacttgc actgagagtt gaggcatctg      60
ggctgtcnac aggtgtctga cgtgtaaact tggaaatcaaa ctgacttaca tcctttcag      120
attgcaacag aggtttaaag gggggctcca ctttcgagc cagaaggcttccctcaggtaa      180
tgtgtctaaa gaatggatga gcttgaactt ctccagcgtc cccaggacca gctcccagac      240
gagaagcgcg atttcttttc agcagcttt taagcagatc tctggcttct tgnngtgaggt      300
agggaggcaa attgag      316

<210> 100
<211> 425
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(425)
<223> n = A,T,C or G

<400> 100
accgctttca gaaagtttat atgggttatt ctccagcctc tcttttatgc ctggaccc      60
ctgtttatca accccaaacc aattacgtat ctggaaagttt tcaataccgt ggcacaggc      120
actttgaca tttaattta ttacttttg ggaattaaat ccttagtcta catgttggca      180
gcattttac ttggcctggg ttgcacccca atttctggac attttatagc tgagcattac      240
atgttcttaa agggncatga aacttactca tattatggc ctctgaattt acttaccc      300
aatgtgggtt atcataatga acatcatgat ttccccaaca ttccctggaaa aagtcttcca      360
ctggtgagga aaatagcgc tgaatactat gacaacctgc ctcaactacaa ttctggata      420
aaagg      425

<210> 101
<211> 156
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(156)
<223> n = A,T,C or G

<400> 101
actgacttgg gaatgtcaaa attcttattt atgatcttcc gagtggttgc ctgagctttg      60
ttggccctca actgcaggca gagaaccagg agcagggtgg cagggtggc cctgaacagg      120
agctggagca agcgcattgtc ngagaaaaca gaaggc      156

<210> 102
<211> 230
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(230)
<223> n = A,T,C or G

<400> 102

```

actccaggcc	gggnctcagg	ttatcaaag	tgcaggagct	ctgatcagca	tggaccactt	60
cttccaaaga	atttccctgc	tggccgttg	taggggttgc	ggttaattctca	taaccagtaa	120
tgtctgggt	ggtgctcctc	tcccaggaga	ctgtgagcac	tccagtgta	gggtttgcct	180
ccagatgcaa	gnntngtngt	ggagacaatg	gtgnaccac	tttgtnnaca		230
<210> 103						
<211> 404						
<212> DNA						
<213> Homo sapien						
<220>						
<221> misc_feature						
<222> (1)...(404)						
<223> n = A,T,C or G						
<400> 103						
actgtgaacc	ctgnggnttc	nangcacct	acctggagct	ggccagtgt	gtgaaggagc	60
agtatccggg	catcgagatc	gagtcgcgc	tcgggggcac	aggtgcctt	gagatagaga	120
taaatggaca	gctgggtgtc	tccaaagctgg	agaatggggg	ctttccctat	gagaaaagatc	180
tcattgaggc	catccgaaga	gccagtaatg	gagaaaaccc	agaaaaagatc	accaacagcc	240
gtcctccctg	cgtcatacctg	tgactgcaca	ggactctggg	ttctgtct	gttctgggt	300
ccaaaccttg	gtctccctt	gtctcgctg	ggagctcccc	ctgcctctt	ccccacttta	360
gctccttagc	aaagagaccc	tggcctccac	tttgccttt	gggt		404
<210> 104						
<211> 404						
<212> DNA						
<213> Homo sapien						
<220>						
<221> misc_feature						
<222> (1)...(404)						
<223> n = A,T,C or G						
<400> 104						
accaggttat	ataatagtat	aacactgcca	aggagcggat	tatctcatct	tcatcctgt	60
attccagtgt	ttgtcacgtg	gttgtgaat	aaatgaataa	agaatgagaa	aaccagaagc	120
tctgatacat	aatcataatg	ataattattt	caatgcacaa	ctacgggtgg	tgctgaacta	180
gaatctatat	tttctgaaac	tggctcctct	aggatctact	aatgatttaa	atctaaaaga	240
tgaagtttagt	aaagcatcag	aaaaaaaaagt	gggtattct	acaagtcaagg	acattctacg	300
tgactataat	ataatctcac	agaaaatttaa	cattaatacn	ttctaagatt	taattcttag	360
antctngtta	aacaaagttag	ctcctgtgga	natgattggc	atca		404
<210> 105						
<211> 325						
<212> DNA						
<213> Homo sapien						
<220>						
<221> misc_feature						
<222> (1)...(325)						
<223> n = A,T,C or G						
<400> 105						
acagcagaag	ccagtctang	atggtgtgat	tcaattctg	cctctagttat	ttctttgtct	60
tgtttttct	tcaattttaga	agtgagcatt	gtgttctcag	ctatcagaac	tttaagctgc	120
ccactatatt	gagatgccct	tttagctaat	gatctcttt	tcaagttttag	ggtcatactga	180
agttcagcat	tctttctt	taaaatctta	atgcctcaa	agtattttatt	ttcctttcc	240

tggtattggn	gtttcagngt	ggcttattcc	agtttttagca	tggcaattnc	ctttttcaac	300
atgcaatttt	catgtaaagag	ataat				325
<210> 106						
<211> 444						
<212> DNA						
<213> Homo sapien						
<220>						
<221> misc_feature						
<222> (1)...(444)						
<223> n = A,T,C or G						
<400> 106						
actgtcttca	atnstatcg	tgcagggtgc	taccacaggo	aaacagttt	ctccccattt	60
tgttagtaatg	tgattttcct	attagcaaaa	agaggtcacc	agccccctgt	gacttaaggg	120
actcaagtca	caggatgggg	atttccttct	aatatttttt	atttngttgt	ttgaactctt	180
gatgcaacat	tgttagagcag	ggtgttcagg	acctgctgtg	cccaaggac	tgataaaagga	240
aaaagctcta	tttattcttt	ttgtgattt	atgcacagat	gaaaaactta	acacacaata	300
acagaagttg	gncttataa	aatcacatcc	taggcttca	gcttncgt	aagcagacga	360
catcttcagt	tttctagctc	ttgnagnn	aacacngnaa	catcaatgtat	gcataatgtnc	420
agaatcagtt	acaaagacca	tccg				444
<210> 107						
<211> 287						
<212> DNA						
<213> Homo sapien						
<220>						
<221> misc_feature						
<222> (1)...(287)						
<223> n = A,T,C or G						
<400> 107						
acctgcactc	gnacntcagg	cantaggcct	ccacgtcatg	gccaggca	ggcatggct	60
ccaccacgtg	caggcagt	cagtccttct	gggatacatt	ctggttgt	atgtgcccac	120
tgtatgtttct	ataaggtggg	acagatgtat	ttgcacccga	tatcttca	actcttgg	180
gctncagctg	ggggcacca	caaaccaccc	accacagcca	ccaaagataa	nagcttcatg	240
cttacatcangc	ttgctggcc	agnaaagcc	gacacctaca	agcccn		287
<210> 108						
<211> 478						
<212> DNA						
<213> Homo sapien						
<400> 108						
acatgtgcaa	gaatttggaa	aaggcaggca	ttttccctca	tctctccat	agggaaatatc	60
acagcatctg	tctctactgg	tccacactgg	actgcagaca	atgtcaaaac	tctggattt	120
gaatgcggct	gatttcctt	cccctttaag	gagtttcca	agaatttcat	aaccatcagt	180
tgttatattt	ccagcttct	tgatgtctt	ttctataatt	tcatagcagt	caatgtaaat	240
cttaacactt	tttggaggta	ctacaatatg	aaccttgtg	aaacttccat	aaaataatgt	300
ctttacttct	tctgtgtcaa	atgtacatgt	ttgcaccc	ccttggat	ccttggtaaa	360
gaatgataac	gtcttgctag	aaggatctgc	aatcaactcca	acttgggtt	tgtatgtct	420
gtctgtgatt	tgccaaattt	caaaagggtc	actgggagtt	tctggagaa	gtctgaat	478
<210> 109						
<211> 361						
<212> DNA						

<213> Homo sapien
 <220>
 <221> misc_feature
 <222> (1)...(361)
 <223> n = A,T,C or G
 <400> 109
 gaatttttct tctanaataa gtattctgtt gacacagact attggtaaga ttttcaacat 60
 aaggtaatgc taggactggc ctcctagcat gagttgtgag taaagatctg gtctgttgtt 120
 tctccaaaag aagnntctta ctgcgttgc tctcatgagtt ttctgttct gcttcttctt 180
 tttcatattg atataacgg nttttaaat ggttattgtt attaaatatc tcctcatttt 240
 tctcttttag gagatgtatg tgcattttcc tctcaagaaa atgaatatca attgttatct 300
 tgcttttgtt gncagcttc ttatgtgcat gaactaattt ctgttgaagc cacatatttt 360
 t 361
 <210> 110
 <211> 305
 <212> DNA
 <213> Homo sapien
 <220>
 <221> misc_feature
 <222> (1)...(305)
 <223> n = A,T,C or G
 <400> 110
 acataatgac tnncanagt aagctgattt gctgcgggtt tggagtaaat ataaagcttc 60
 cgttcctggg aatccgcact acttggatca cgtgcctggc ctaccaaattt ctgcggaaaa 120
 ctatgtgcct tatcccaccc tnnaatctgn ctccctattt ntcagctgtt ggatcagaca 180
 atgacattcc tntagatntg gcgatcaaggc attcacanacc tgngccaact gcaaacggtg 240
 cctncaagga gaaaacgaag gcncaccaa atgnaaaaaa tgaangnccc ttgaatgtac 300
 taaaa 305
 <210> 111
 <211> 371
 <212> DNA
 <213> Homo sapien
 <220>
 <221> misc_feature
 <222> (1)...(371)
 <223> n = A,T,C or G
 <400> 111
 cggggggccag cggggggat tcagccatcg atcaaactca aaacctggaa tgatatccac 60
 tctcttttc ttaagctca gggaaatattt caagtagaaag tccagaaagt catcggtt 120
 gatgttcgg aatttgaattt catgcacata ggccttggaa aaactgtcaa actgtatcc 180
 atcacccacc aagtggggca ggtatgagac aaagcagaaaa ctttctgtt agggggtctc 240
 attataggtt tcgtccgggtt caacgcgtt ttcaatcttc acgcggagct tggtgagtgg 300
 gttttcttcc caagtatgtt ccatgtgctt acgcagcaga ncccgccccg ttgcagccctc 360
 caagcagggng t 371
 <210> 112
 <211> 460
 <212> DNA
 <213> Homo sapien

```

<220>
<221> misc_feature
<222> (1)...(460)
<223> n = A,T,C or G

<400> 112
acatcttagg ttttnttcc tttantgtga agaggcggtt ccacccaaccc acagctctgc      60
gtcgagttt tactagattt ctgcaattt catggaatct ttgctgttgt tcagtggtcc      120
atttatttggaa gccaaaaattt ctaggcgct agaatggaa caaggtatgc agccaagcac      180
aaaaacataaa caaaacaggaa aacgcggac agaacagatg gatctagata gtagataatc      240
agaaacacca aagaaaccac acccatatgc gcaagggtggaa accaggttct ttctcatcg      300
aggactttat cagccatcaag catcaattt ccccatcctt gcagctgttc ttccagactt      360
gcagtcctcg cagccagcaag gttgggtgtc gcgattacct ccctccgcca tcgtctcg      420
gatgcagtctt ctacaagcgc aggccacctc cccaaacgagt      460

<210> 113
<211> 204
<212> DNA
<213> Homo sapien

<400> 113
gagaagacag cagagctgct ttccgcctct ttgagaccaa gatcacccaa gtcctgcact      60
tcaccaagga tgtcaaggcc gctgctaattc agatgcgaa cttctgggtt cgagcctct      120
ggcgcccttag cttggaaacctt gggaaagaat atttgcattt gggcttagat gggccaccc      180
atgacctcga gggacaccccc cagt      204

<210> 114
<211> 137
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(137)
<223> n = A,T,C or G

<400> 114
accgcaagaa atgggacagc aacgtcattt agacttttttga catcgncgc tngacagtca      60
acgctgacgt gggctatttttac ttctggaggt gtcccaagcc cctgaagaac cgtgatgtca      120
tcacccctccg ntcccttg      137

<210> 115
<211> 278
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(278)
<223> n = A,T,C or G

<400> 115
gcggggcggt ttntggactt gctcatttttac agagcatgctt tggctttcac ccttggcatg      60
ttctccgccc gcctctcgaa cctcaggcac atgcgaatgtt cccggaggtgt ggacaacgctc      120
cagntcctgc cctttctcac cacggangtc aacaacctgg gctggctgan ttatggggct      180
ttgaaggag acgggatccatcgatcanc aacacagtgg gtgctgcgtc tcanaccctg      240
tatatcttg gcatatctgc attactgccc tcggaaagc      278

```

```

<210> 116
<211> 178
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(178)
<223> n = A,T,C or G

<400> 116
acaccgtcat angtcaaaag tncagtgc tgccatcttgc atcaaatgtt cttaaggcag      60
tgactggcta tcaaccacag nttctgtctc cccagntgca aacacaggat ccatgcaaca      120
gttctgagac catacactta gaaaccacng ggagatgcgg atcanatgca naactnnc      178

<210> 117
<211> 360
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(360)
<223> n = A,T,C or G

<400> 117
actccccat ggnnggattta ttactattaa agaaaccagg gaaaatatta attttaatat      60
tataacaacc tgaaaataat ggaaaagagg ttttgaatt tttttttaa ataaacacct      120
tcttaagtgc atgagatggt ttgatggttt gctgcattaa aggtatttgg gcaaacaaaa      180
ttggaggggca agtgaactgca gttttgagaa tcagtttga ccttgcattat tttttgttc      240
cactgtggaa ataaatgttt gtaataaagt gtaataaaaaa tccctttgca ttcttctgg      300
accttaaatg gtagaggaaa aggctcgta gccattttgtt tctttgctg gttatagttg      360

<210> 118
<211> 125
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(125)
<223> n = A,T,C or G

<400> 118
gcgtcgtgct atgaccggac ttngtcttga aaggggatga cagcatggga ggcaatggnt      60
ncacatgtaa accccacact gaaagacaag gcactctctc cacagcagcc ccaacaacta      120
gccct      125

<210> 119
<211> 490
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(490)
<223> n = A,T,C or G

```

<400> 119
 nacaaagaaa agcaaaaaga atttacgaag attgtgatct cttattaaat caattgttac 60
 tgatcatgaa tgtagttag aaaatgttag gtttaactt aaanaaaatn gtattngat 120
 tttcaatntt atgttgaat cngngtaata tcctgangtt ntttcccc cagaagataa 180
 agaggataga caacctctta aaatatttt acaatttaat gaaaaaaagn taaaattct 240
 caatacnaat caaacaattt aaatattttt agaaaaaaagg aaaagttagat agtgatactg 300
 agggtaaaaaaaa aaaattgtattt caatttatg gtaaaggaaa cccatgcaat ttacctaga 360
 cagccttaaaat tattgtctgg tttccatctg ctgcatttc agacattttt ttttcctt 420
 actcaatttga taccacaga aatatcaact tctggagttt attanatgtt ttgtcacctt 480
 tctnaagttt 490

<210> 120
 <211> 361
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(361)
 <223> n = A,T,C or G

<400> 120
 caggtacagt aaaattaaca cttccgttac agaaaatgtt tgacgcaaattt aatataaaaat 60
 taaaagggttga aaaaaagggtt acactgtttt cctaagatac aatttactt ttacaaccag 120
 ggtccacagg tccaggctgc anagcgggca tcaggaagca gagectncca cctgcttctg 180
 ggggacctgg taataaaaat cagcccatga tggcgtatg gcctctcaga caccacacgc 240
 tgcctaaaca cctagagctc tggaaatagt caacaggaga gtgatttcca tggggaaat 300
 tttaaaanaag atgcacatgg gacaggcaat agaaaagttt ccaaggntaa atttggtacc 360
 t 361

<210> 121
 <211> 405
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(405)
 <223> n = A,T,C or G

<400> 121
 acacaaaacc ttttnacata ttggggcattt accgctccaa attgtactt atcccttaag 60
 ttcacaatatac agaatttctt caccaattaa gtaataaccc tcattacaaa taaagtgcatt 120
 ctgataacca aactcgtaag tcccatgtc agggactgtc tggccatttt aaggatcccg 180
 tatatatggt catgtttctt tataacaggc gtcattgttgc acaggttagcc atgtatgtt 240
 ccgatcacaa atagtatggg tggcaagagg aggtatataag aagtatctt ttttacactt 300
 ataatctact cgttccaccaa tctcatatgtt gggttttgtt ttaccaatgtt gcctccatattt 360
 cttcaaatgtt tgggtggctt ctcacaggca tcnggcanaa ngagt 405

<210> 122
 <211> 152
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(152)
 <223> n = A,T,C or G

<400> 122
 accccgctcc gttgnacacag atcgctgtct gcccactcca tcggccattc acttggcagg 60
 tgcgattggc agagccccgg agagttaac cgtcatagca gtggaaagag atotcatcac 120
 tcacattgtta ttagggagac cggggccaan ta 152

<210> 123
 <211> 336
 <212> DNA
 <213> Homo sapien

<400> 123
 acatctgaca tatttatata gcacataaat tagggagtgc tctgaccctt gcccgtggag 60
 cccaaagcaact gaggcaggag gtgaacgcca gtccagaaag aaggtgtctgg agcccccgtct 120
 ctgttcttc catcacgggg ctccccctagg gcctcccccag gcctcccttgg ctcaagtccag 180
 gtgtctgcag gaggaaagggtg ttgtctgcat ttagtgtctg agactgggtt tgaggaggca 240
 ccagataaaa ggagatacac ttgcagctt aaagtcaagct tcaaacccca gggcttgtaa 300
 ttccaaagagg agggtgggga ggcgaggcca tagtct 336

<210> 124
 <211> 253
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(253)
 <223> n = A,T,C or G

<400> 124
 ctgcaagagc ccagatcacc cattccgggt tcaactccccg cctcccaag tcagcagtcc 60
 tagccccaaa ccagcccaga gcagggtctc tctaaagggg acttgagggc ctgagcagga 120
 aagactggcc ctctagcttc taccctttgt ccctgttagcc tatacagttt agaatattta 180
 tttgttaatt ttattaaaat gctttaaaaa aacaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 240
 aaaaaaaaaaaaaaagntt gtn 253

<210> 125
 <211> 522
 <212> DNA
 <213> Homo sapien

<400> 125
 acaactgcaa gtctaagata atgttcattc attccatca taaaatgttaac attctaaata 60
 ggtgttcttctt gatgtcatct gtcagaattt cttttaact ttttcttcat cttaacattt 120
 atcaaaagttc atccatttattt ctcttgcctt gatttcgag agtttccaaat ttttcaactt 180
 ttaaggcagc gattgtttt gcatctctgg tattttatctg ctcttcttga aaatttctct 240
 ttgcttcttc gttagaaataa aacttaacag ttggataggc cctgatccca gctttctggc 300
 atgtctgagc ataaggctga cagtctactt ttccagcttt cacttttctt ttaatcatcc 360
 tagccaaagag ctcaaaattct ggagcaaaaat tctggcaagg tccacaccaa ggagcataga 420
 aatcaatcac ccaatgattt ttcccttgcata gaacttttctt actgaaagtc tgaggtgtta 480
 gatctgtgga tacttgaggt aaaaatccca gaccccgatgc tc 522

<210> 126
 <211> 374
 <212> DNA
 <213> Homo sapien

<220>

```

<221> misc_feature
<222> (1)...(374)
<223> n = A,T,C or G

<400> 126
tttttaagat attaacttta cttttataaa tccttgcgtg aatgaaaaaa aaaaatcaag      60
gcatacaaat ttcattgtgt tctacattt taaataccat cctttgtctc cgtaaaaga      120
tttcatcca ttatccaaa aacctttaa gtcaactgt ccaatttaag acagagtcaa      180
gacattttg agtatctgaa ctaagcattg tcctgactga aacgaagtaa gaactcaatg      240
agagtccttgc tgggcctccc aggcatgcct ttccgttagat agggaaacttc atctttgttg      300
gncatcacgc ctgttatgtc taaaatgtgcc cacttaggat gagttacgaa ttcttcagg      360
aatgtcagc ctgt      374

<210> 127
<211> 130
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(130)
<223> n = A,T,C or G

<400> 127
aaagccaaga cngccattgg cactgctatg gtaaggncac agggcancca gggcctctg      60
gcaaaaggng atacnaccag cactatnaac agacaggaca tggttgagag gnagnctaca      120
caantcctaa      130

<210> 128
<211> 350
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(350)
<223> n = A,T,C or G

<400> 128
acactgattt ccgntnaaaa gaancatcat cttaaccttg acttttcagg gaattactga      60
actttttctt cagaagataa ggcacagcca ttgccttggc ctcacttcaa gggctgtcat      120
ttgggtcctc tggctctttt ccaagnttcc cagccactcg agggagaaat atcgggaggt      180
ttgacttctt ccggggcttt cccgagggtt tcaccgttag ccctgcggcc ctcagggtg      240
caatccttggaa ttcaatgtct gaaaccttgc tctctgcctg ctggacttctt gaggccgtca      300
ctgccacttctt gtcctccagc tctgacagct cctcatctgt ggctgttga      350

<210> 129
<211> 505
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(505)
<223> n = A,T,C or G

<400> 129
acaataccaa agttcataaa tgctaaagaa aacccaaaaca aaagacaatg gtttacacag      60

```

gaaaataacc ctaaggcaat atgaaaacag tcataattta ttactgataa agagtaaagg	120
catccttccc atagaggggg ggaattcaca gggAACACTA attatatcg atgaaccacg	180
gggatagaaa ataggccat tttaaaatt cattgagaaa ttattactt ttctccacaa	240
ctgtgattct atacaaaata taaACCTGC aaACCTTATG tgctacctga cagataaaag	300
tagcaggagc cagactctg aagcacttga gactgattt tacaaagtcc aggaagagca	360
atgattccag tgcgcgtgc tgatcatgt gtgagcctaa catgttattc agctctgg	420
gcagccccat ctacatgggg cccagttgt tttagggag tcacagatta ngcaggcaac	480
cgaggggcat gattaaaaaa gcaca	505
<210> 130	
<211> 526	
<212> DNA	
<213> Homo sapien	
<400> 130	
acaaaagagc ctgattctt ttaattccac aaatacctag catctcaaag taacatgtaa	60
acaaacttct atgtgcgtca atgaatcctt ccaatttgcg taataaacta aatagtattg	120
gatcttagtat atgactttca tggtaagtt atggttctat ccattactt aacaatatta	180
ctgtatgtac agagaaaaat ttcaactat tgacttatt taaaacaaac tgacaagtcc	240
aagcacctgt ctgcagaaaa ggcagcagca tttttttttt tttaacatc tcaaagttaag	300
atttggccta agcccttaat acctttctga acagccatgc aactaaacac cctcaggaga	360
tgttacataa gggagagaag aacatggagc aatttgcact ttccccta gataatatta	420
acaaggtaaa gcaaatccag atctttatga atgaatggct gtcatgttta atacacttgg	480
agctctataa aactagagcc actatcatat atgttataat agat	526
<210> 131	
<211> 477	
<212> DNA	
<213> Homo sapien	
<400> 131	
ctcagttttc ccagcaacag atgcctctga gcaattttt agtcaagtga cggcgtgaa	60
atactttct cattacatgg aggagaacct catggatggt ggagatctgc ctgtgttac	120
tgtatctca agacctcgcc tctaccctct tcagtggcta aaatctgata aggcctaat	180
gatgtctttt aatgtatggca ctttcaggt gaatttctac catgatcata caaaatcat	240
catctgtac caaaatgaag aataccctt caccctacatc aatgaggata ggatatctac	300
aactttcagg ctgacaactc tgctgtatgc tggctgttca tcagaatcaa aaaattgaat	360
ggaatatgcc ctgaacatgc tcttacaaag atgtacttgc aagactttc gaatggaccc	420
tatggactc ctctttccca ctgtgagatc tacagggac caaaaagaat gatctag	477
<210> 132	
<211> 404	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(404)	
<223> n = A,T,C or G	
<400> 132	
accacacgan cgggnatcnt ttgnacatag tgagacccgg ctgattccca tacatgaatc	60
cattcatgga gtgcattttt ttagatnct gaaagtctt atcttcctt tccacctgtat	120
cagggcagt tgtaaacatn cctaataattt tcttcaggaa gtaaaactctc attctcatca	180
aataactgtat gaaaacaaata gaattccctt tctacatctt tctgtctccc atttgcataat	240
aaacttcctt tcttgcataat tttcatggc ccaataagcc cagtgaatat atctttatgt	300
ggatccacag cagaataata catcttagct agacacacag ggatctgcatacgnggg	360
ctacttctttt ggggacagcc cttcatacgn gaatgtttnt gtgg	404

```

<210> 133
<211> 552
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(552)
<223> n = A,T,C or G

<400> 133
accccaaatt atctctctcc tgaagtccctc aacaaacaag gacatggctg tgaatcagac 60
atttggggcc tgggctgtgt aatgtataca atgttactag ggagggcccccc atttggaaact 120
acaaatctca aagaaaactta taggtgcata agggaaagcaa ggtataacaat gccgtccctca 180
ttgctggctc ctgccaagca cttattgtct agtatgttgc cccaaaaccc agaggatcg 240
cccagttgg atgacatcat tcgacatgac tttttttgc agggcttcac tccggacaga 300
ctgtcttcta gctgttgtca tacagttcca gatttccact tatcaagccc agctaagaat 360
ttcttaaga aagcagctgc tgctttttt ggtggcaaaa aagacaaaagc aagatataatt 420
gacacacata atagagtgtc taaagaagat gaagacatct acaagcttag gcatgatttg 480
aaaaagactt caataactca gcaacccagc aaacacaggg acagatgang agctccacca 540
cctaccacca ca 552

<210> 134
<211> 496
<212> DNA
<213> Homo sapien

<400> 134
acattgatgg gctggagagc aggggtggcag cctgttctgc acagaacccaa gaattacaga 60
aaaaagttcca ggagctggag aggcacaaca ttccttgggt agctcagctc cgccagctgc 120
agacgctaat tgctcaaact tccaacaaag ctgcccagac cagcacttgc gttttgattc 180
ttctttttc cctggctctc atcatcttc ccagcttcag tccattccag agtcgaccag 240
aagctgggtc tgaggattac cagcctcagc gactgtttc cagaataatc ctgaccac 300
aggacgttaac agaaaaatctg gagacccaaag tggtagatgc cagactgacg gagccacctg 360
gagccaagga tgcaaatggc tcaacaagga cactgcttga gaagatggga gggaaagccaa 420
gaccctggc ggcacatccgg tccgtgtgc atgcagatga gatgtgagct ggaacac 480
ttttctggc cacttt 496

<210> 135
<211> 560
<212> DNA
<213> Homo sapien

<400> 135
actggggagtgt atcactaaca ccatacgtaat gtctaatatt cacaggcaga tctgcttggg 60
aaagcttagtt atgtgaaagg caaatagagt catacagtag ctcaaaaggc aaccataatt 120
ctctttgggt caggtcttgg gagcgtgatc tagattacac tgcaccattc ccaagttaat 180
ccccctgaaaaa cttactctca actggagcaa atgaactttg gtcccaaata tccatcttt 240
cagtagcggtt aattatgtct tgtttccaaac tgcatttcct ttccaaattga attaaagtgt 300
ggcctcggtt ttatgtcattt aaaattgttt tctaaatgtat tgctgcctctt attatggcac 360
ttcaattttt cactgtctt tgagattcaa gaaaaatttc tattttttt tttgcatcca 420
atttgccctg aactttaaa atatgtaaat gctgcccattt tccaaacccca tcgtcaagtg 480
tgtgtgtta gagctgtgca ccctagaaac aacatattgc ccatgagcag gtgcctgaac 540
acagacccct ttgcattcac 560

<210> 136
<211> 424

```

```

<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(424)
<223> n = A,T,C or G

<400> 136
accagcaaat ctccattagc atttctcagg tttcatgatc cttttcagat atgttgggtt 60
attttatgtatatattgtttt agaaacaaaaa atccacctga tattaaaaca aaccaaaaaa 120
aatcataaaa gcaagcaaat gaacaaaaaa cccttagttt gttgtgtttt tctttcacat 180
ttccctacagg gagatgttataatctcagat actttcaaaa tctaataatgtt aagtaaaatt 240
agtgccttaa ccaaacagta agataccaaa gaatcccca tcacaagttt ctgaatcaaa 300
cttctcatgatatttgcgtt atattcagat ttgaagattt tttaaatttta gaatttaaaa 360
caaacttttag actgtcgatt ttccatattt caaagactgt agctgtntgc agcatataaa 420
tgga 424

<210> 137
<211> 392
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(392)
<223> n = A,T,C or G

<400> 137
tgcggggntg aaggctagca aaccgagcga tcatgtcgca caaacaatt tactattcg 60
acaatatacgacgacgaggag tttgagtatc gacatgtcat gctgccaag gacatagcca 120
agctgggccc taaaaccat ctgatgtctg aatctgaatg gaggaatctt ggcgatcagc 180
anagtcaaggatgggtccat tataatgatcc atgaaccaga acctcacatc ttgtgttcc 240
ggcggccact acccaagaaa ccaaagaaat gaagctggca agctactttt canccctcaag 300
cttacacag ctgncccttac ttccatacat ctttctgata acattattat gctgccttcc 360
tgttctcaactctganatntaaagatgttc aa 392

<210> 138
<211> 284
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(284)
<223> n = A,T,C or G

<400> 138
tgcctgtca cctctttgtctgaaatatgg caagacttgg aaaaatgttt gccccttagaa 60
tctatctcac tacttagttt agttgtctcc tttggggctgttgcacagttc tggccctgtat 120
ctggaaacaga ctcccttttcaaaactgaa cttgaccaca tcaaaagntt gnaaaacaat 180
ctccatggta attaaacttg cattcaacac catatggnaa cagaagatgg caggaggata 240
anatncagat cttatgtatcttccangnan ggcatgttac atga 284

<210> 139
<211> 249
<212> DNA

```

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(249)

<223> n = A,T,C or G

<400> 139

gaggaagggg ggactgaatc tancaccntg acngaactag agacagccat gggcatgatc 60
atagachnct ttacccgata ntcgggcagc gagggcagca cgcagaccc gaccaagggg 120
gagctcaagg ggctgatgga gaaggagcta ccaggcttcc ngcagagngg aaaanacaag 180
gangccgtgg ataaattgct caaggaccta gacgcccnatg gaggatgccc aggtggactc 240
cagcgagnt 249

<210> 140

<211> 390

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(390)

<223> n = A,T,C or G

<400> 140

tcataatggc tggggcagct ataatnnact acaanaatca natgtttcac atctagacct 60
cgggcagcaa cagaggttagc cacaagaagt ttgcangtcc cattctaaa gtcatttatg 120
atgctatctc tgtcatattg atcaatgcct ccatgaagag acatgcaagg ataagatgct 180
ctcattaaat ccttaagaag accatcagca tggctctgct tatccacaaa tataatgaca 240
gatcctgact ctgtataatg gcctagaagc tcaagtaact tcaagaattt cttttcttct 300
tcaatcacaatcacttgc tccacatct gagcaaacca cactctgcc tccaacttgt 360
acctgccccg ggcggcgct caagggcgaa 390

<210> 141

<211> 420

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(420)

<223> n = A,T,C or G

<400> 141

gacactcagg gaaaagcatn ngncaaanag agcttaaat gcatcgccaa cggggtcacc 60
tccaaaggct tcctcgccat tcggagggtgc tccactttcc aaaggatgat tgctgaggtg 120
caggaagagt gctacagcaa gctgaatgtg cgcancatcg ccaagcggaa cccngaagcc 180
atcaactgagg tcgtgcagct gccaatcac ttctccaaca natactataa cagacttgnn 240
cgaagcctgc tggaatgnga tgaanacaca gggcagcaca atcaggagac agcctgatgg 300
anaaaantgg gcctancatg gccaggcctc ttccacatcc tngcangaca gaccactgtg 360
cccaaacaca cccnctgagc tgacttnac aggagacgca cnaaggagcc cggcagangc 420

<210> 142

<211> 371

<212> DNA

<213> Homo sapiens

<400> 142

gggttcgaca atgctgatcc gcaattagaa gacactggta agctgtgtta cactgggctt 60
 cattgaaatc ttcaaggata tagccagctc ctgctcgaaag ctgggattct gtatactgt 120
 tggtaagg aggaatttcc aaaaatttcc ccttcttc actgcttctt gtaggaccat 180
 ctggcagttt ggagcggctg gccaacttgt cactgggtgt ggccatggta aggagaaaatg 240
 cgtagcccag aaacaagggtc ttgttggagag gcaaaggccc tctctgtct tccaggcag 300
 agggttcacc ggtgtgtct ccactctcac aggggctcac aaactctctt gccccactt 360
 gcaccaggtt t 371

<210> 143
 <211> 270
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(270)
 <223> n = A,T,C or G

<400> 143
 ggtggctgtg atnacctttt ttagttaca aataaaaaaag ntaaaaaagaa atactgtgtt 60
 tagggtaagg taacannttc atctaattc agggaggtga agangaggcn ctgccttcta 120
 ggnngctgtga ccttcttc cttcngattc ttccnccacct tgggnaacat ttccccgt 180
 atgctggaan tacttcggng ttctgcggtg gccatgntga acatctgtat aactgaaant 240
 ncatccnaat gcacacgaag anatagnna 270

<210> 144
 <211> 259
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(259)
 <223> n = A,T,C or G

<400> 144
 ttctctttgc tttttataat tttaaagnaa ataacacatt taactgtatt taagtctgtg 60
 caaataatcc ttccagaagaa atatccaaga ttctgtttgc agaggtcatt ttgtctctca 120
 aagatgatta aatgagtttgc ttccagata aagtgtctt gtccagnaga actcaaaagg 180
 cttcaagct gttcagtaag tggtagttca gataagactc cgnccatacga attccagctt 240
 cccgtgcccc ctgtaccc 259

<210> 145
 <211> 433
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(433)
 <223> n = A,T,C or G

<400> 145
 accacatnta ccatacgta attagttta atttcacat gaatcaaagg ttcccttca 60
 tgtctattta cagtccaaatt gtgc当地act cttacttgtg tgctgactaa caaggcattt 120
 aggtgtgcag catccttagag tgctccaggc cagtgctcagc gttctcgagg gtaaaagggtg 180
 ccacttggta gcaatgatat tccagaatta aatgggtttt tgtagccatg gagactgcat 240
 ttatataaat gtagcctgtta gcttaaggta actaaaccta atgctgtgt taaaaacagt 300

ttatTTtaat attaaaatac agttgattag caacagcggt gctgtatTTt aagagacact 360
ttattGGAAG tgcaatcata gttattgtt ttcaacaaTTt tacagngcat tctaattact 420
gatgggtgca att 433

<210> 146
<211> 576
<212> DNA
<213> Homo sapiens

<400> 146
acctcaggcc tttgcaccc ttgtctgaa atatggcaag acttggaaaa atgtttgccc 60
ttagaatcta ttcactact ttgttagtt gtctccttg ggcctggca cagttctggc 120
cctgatctgg aacagactcc ctttctaaa actggaccc ttttccatca aaagtttgc 180
aaacaatctc catggtaatt aaacttgcatt tcaacaccat atggtaacac aagatggcaa 240
aggataagat tcagatctt gatcttcca agtagggcat gtttagatgtt agaaggatta 300
gttgcagact ggtatctggc tcaggcttgg gcatgaaggaa aactgtctcc catgtggttt 360
ggaagagttt ggggctccct gagcttattt gtgaactata cgggttccat ccaaggaatg 420
gtatgtatgtt ggcataaaaaac cattcttag acaactgaag atggtccccct tctgttagcca 480
gaaacactag ctgtcctgca ttgccatttc ctttacccca ggcggcctgc agaaggaaag 540
gccataatta attaaaaggc ttaatgaagt tttgga 576

<210> 147
<211> 300
<212> DNA
<213> Homo sapiens

<400> 147
ccagccccca ggaggaaggt gggctctgaat ctgcaccat gacggacta gagacagcc 60
tgggcatgtt catagacgtc ttaccggat attcgggcag cgagggcagc acgcagaccc 120
tgaccaaggg ggagctcaag gtgttatgg agaaaggagc taccaggctt ctgcagagt 180
gaaaagacaa ggatgccgtg gataaatttc tcaaggaccc agacgccaat ggagatgccc 240
aggtggactt cagttagttt atcgtgttcc tggctgcaat cacgtctgcc tgtcacaagt 300

<210> 148
<211> 371
<212> DNA
<213> Homo sapiens

<400> 148
acataatcct cataatggtt gggcagcta taatTTcta caagaatcag atgtttcaca 60
tctagaccc gggcagcaac agaggtagcc acaagaagtt tgcaggccc attcttaag 120
tcattttatgtt tgcattatgtt tcaaatggcc tccatgaaga gacatgcaag 180
gataagatgc ttcattttttt tccatggaa gaccatcagc atgttccctgc ttatccacaa 240
atataatgac agatcctgac tcttgataat ggccttggaa gtcacgttac ttcaagaattt 300
tcttttccatc tcaatcaca atcacttgc tgcctcacatc tgagcaaacc acactcctgc 360
ctccaaacttg t 371

<210> 149
<211> 585
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(585)
<223> n=A,T,C or G

<400> 149
 cgaggtacan cactgctaaa tttgacactn angaaaaagc attcgtaaa gagagcttaa 60
 aatgcacatcgcaaacgggtc acctccaagg ttttcctcgc cattcgagg tgctccactt 120
 tccaaaggat gattgctgag gtgcaggaag agtgctacag caagctgaat gtgtcagca 180
 tcgccaagcg gaaccctgaa gccatcaactg aggtcgctcca gctgcccaat cacttctcca 240
 acagatacta taacagactt gtccgaagcc tgctgaaatg tgatgaagac acagtcagca 300
 caatcagaga cagcctgtat gagaatttgc ggcctaacat ggccagcctc ttccacatcc 360
 tgcagacaga ccactgtgcc caaacacacc cacgagctga cttcaacagg agacgcacca 420
 atgagccgca gaagctgaaa gtcctctca ggaacctccg aggtgaggag gacttccct 480
 cccacatcaa acgcacatcc catgagagtg cataaccagg gagaggntat tcacaacctc 540
 ccaaactagt atcatttttag gggngttga cacaccagtt tttag 585

<210> 150
 <211> 642
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(642)
 <223> n=A,T,C or G

<400> 150
 acttncgggt tcgacaatgc tgatccgcaa ttagaagaca ctggtaagct gtgttacact 60
 gggcttcatt gaaatcttca aggatatacg cagctctgc tcgaagctgg gattctgtat 120
 actgcttgcatt gaaaggagga atttccaaaa attccttcctc ttcttcactg cttccctgtat 180
 gaccatctgg cagtttggag cggctggcc acttgcact gttgtggcc atggtaagga 240
 gaaatgcgta gcccagaaac aaggtttgt tgagaggca agggcctctc tgctcttcca 300
 gggcagaggg ttccacccgggtt ttgtctccac tctcacaggg gtcacaaac ttcctggccc 360
 ctactgcacc aggtttact gtggcagact tgcacactcg cttggcaggg gaccgttccct 420
 cttcagaagt gataagttt ctttgcctg agagaactcc catggagggca cgaggacttt 480
 ctgtgatctt tccggtaggg gttgtgtgc tactggagggc agtangggtg gctggggagc 540
 tgacgtaact ggcgcgttcc cgcttccttc caccaaattt ctaagctgat atctgctgcc 600
 tttgtaaagaa gnggtactgc ttcatanggg ccaagcccat ac 642

<210> 151
 <211> 322
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(322)
 <223> n=A,T,C or G

<400> 151
 nttggacaac atcttccccgtt ctagtgcggg attacttcgg ttttcgtcggtt tggccatgg 60
 gaacatctga tgaactgaaa ttccatcgga atgcacagga agatatacgat gatcttcaaa 120
 aatgtccctt ccaggaccac catactgggg aagtttttc ggggtgcctgc naatgggctg 180
 caccctgggg ctggggcccgaa gctctagtc tgcacatgcca tcgccactga aatcggttt 240
 cagatgatta gtctttcat gccccgttcca ttttcgggtt tttctccagttt gttcagaaat 300
 tcaaatacgatt aacttctggg aa 322

<210> 152

<211> 262
<212> DNA
<213> Homo sapiens

<400> 152
acaaagtctt ctctttgctt tttataattt taaagcaaat aacacattt actgtattt 60
agtctgtc aataatcctt cagaagaaat atccaagatt ctgtttgcag aggtcatttt 120
gtctctcaaa gatgattaaa tgagtttgc tttagaataa agtgctcctg tccagcagaa 180
ctcaaaaggc ctcaagctg ttcaagactg gtatcaga taagactccg tcatacgaat 240
tccagcttcc cgtccccact gt 262

<210> 153
<211> 284
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)
<223> n=A,T,C or G

<400> 153
ctcgggagta aaaggtgcca cttggtagca atgatattcc agaattaaat gggttttgt 60
tgccatggag actgcatttataa atgttgcgcctgtactttaactttaactttaatg 120
ctgctgttaa aaacagtttataatattttaataaaatcgttgcatttagcaacagcggtgct 180
gtatTTTaaag agacactttatggaaatgtc aatcatagtttattgttttacacaatTTT 240
ngtgcatctt aattactgtat gggngcaatttactttaatcgnng 284

<210> 154
<211> 531
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(531)
<223> n=A,T,C or G

<400> 154
acccacccta aatttgaact cttatcaaga ggctgtatgg tctgaccatc aaataggata 60
ggatggacct ttttttgagt tcattgtata aacaaatttt ctgattttgga ctttaattccc 120
aaaggattag gtctactctt gtcatttccac tctttcaag ctctgtccac tctaactttt 180
ctccagtgatc atagataggg aattgttcac tgcgtgccta gtctttcttc acttacctgg 240
cctctgtatc aaacagttgc cccttcatt tcataaggtc gaggacttgc gaccctggat 300
ggttctaaat gggaaaagca cggccagatt gtggaaacctg gcttcaacat cagcattctg 360
aaaatattca tcaccatgtat gtctgagatgttccgatgtatgcgtgaacaa atggggaggaa 420
cacattgccc aaaactcactg tctggagatc ttcaacatgt tctccctgtat gaccctggac 480
agcatcatga agtgtgcctt cagccaccag ggcagcatcc agtngacag t 531

<210> 155
<211> 353
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<222> (1)...(353)
<223> n=A,T,C or G

<400> 155

tcttacaagg actgagagag ttacatgttg gaaaaaaaaa agaagcatta acttagtaga 60
actgaaccag gaggcattaag ttctgaaatt ttgaatcatc tctgaaatga agcagggtga 120
gcctgcctc tcatcaatcc gtctgggtgc cagaactcaa ggttcagtgg acacatcccc 180
ctgttagaga ccctcatggg cttaggactt tcatcttaga tagattcaag acctttacct 240
canaatatg taaactgtga ttgtgtttt gaaaaattat tatttctaa accatttaa 300
gtcttgtat atgtgtaaat gatcacaaaa atgtattttt taaaatgttc tgt 353

<210> 156

<211> 169

<212> DNA

<213> Homo sapiens

<400> 156

agtttgttct actacatttt tggtccacta gttcaactttt ctgtgttgat aagcgttacc 60
accaattgca ctttctatacg cctctttac aatgttgctc acttcataaa caacaaaagc 120
agtctcctcc gcagcctggt agtctccat ctttcctccg gcgcgtccc 169

<210> 157

<211> 402

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(402)

<223> n=A,T,C or G

<400> 157

gttaactacc cgctccgaga cgggatttgc gacgagtcct atgaggccat tttcaagccg 60
gtcatgtcca aagtaatggat gatgtccag cctagtgcgg tggcttaca gtgtggctca 120
gactccctat ctggggatcg gtaggntgc ttaatctac tatcaaagga cacccaaagt 180
gtgtgaatt tgcataagagc ttaacctgc ctatgctat gctggggaggc ggtgttaca 240
ccattcgtaa cgttgcggc tgctggacat atgagacagc tggccctgc gatacggaga 300
tccctaatga gttccataac aatgactact ttgaaatctt tggaccagat ttcaagctcc 360
acatcgtcc ttccaacatg actaaccaga acacgaatga gt 402

<210> 158

<211> 546

<212> DNA

<213> Homo sapiens

<400> 158

actttggct ccagacttca ctgtccttag gcattgaaac catcacctgg tttgcattct 60
tcatgactga gtttaactta aaacaaaaat gtaggaaag ctttccttatg ctteggtaa 120
gagacaaatt tgctttgtt gattgggtgg ctgagaaagg cagacaggc ctgattaaag 180
aagacatttgc tcaaccactag ccaccaagtt aagttgtgg acccaaaggat gacggccatg 240
gaaacgtaga tcatcagctc tgctaagttag ttagggaaag aaacatattc aaaccaggct 300
ccaaatggat cctgtggta cagtgaatga ccactcctgc ttatatttc ctgagattgc 360
cgagaataac atggcactta tactgatggg cagatgacca gatgaacatc atcatccaa 420
gaatatggaa ccaccgtgct tgcatcaata gattttccc tgatgttag gcattcctgc 480
catccattgg cacttggctc agcacagttt ggcacacaag gacataatag acaaggccaa 540

aacagt 546

<210> 159
<211> 145
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(145)
<223> n=A,T,C or G

<400> 159
acttttgta taagtttcct aaaaatattt aatactttt ttttcaatt taaattaaat 60
ctnttgatga acaggggggg gntggcaaaa ttccaagcn ctggactgga attttganan 120
aggcatttac ngaccctnat aactt 145

<210> 160
<211> 405
<212> DNA
<213> Homo sapiens

<400> 160
tgtaaatcgc tggttgatt tcctgatttt ataacagggc ggctggtaa tatctcacac 60
agtttaaaaa atcagccctt aatttctcca tgtttacact tcaatctgca ggcttcttaa 120
agtgacagta tcccttaacc tgccaccagt gtcccccctc cggcccccgt cttgtaaaaa 180
ggggaggaga attagccaaa cactgtaagc tttttagaaa aacaaagttt taaacgaaat 240
actgctctgt ccagaggctt taaaactggt gcaattacag caaaaaggaa ttctgttagct 300
ttaacttgc aaccacatct ttttgact tttttataa gcaaaaacgt gccgtttaaa 360
ccactggatc tatctaaatg ccgatttgag ttcgacac tatgt 405

<210> 161
<211> 443
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(443)
<223> n=A,T,C or G

<400> 161
tttgcttta atgaaggaca agggatttaag acncatagag actggccana caaatggaa 60
accgaccaga ccagcccatg accaaaaat cacaggcaga ccacccacaa atgcagaggg 120
ctcagagtcc acagtggcg gttggaaacc agggcccccag ggaatcttc agctgcattc 180
cggctgtat cggcgggcaa caggttagagg tgctggaggg ggctgagtcg tgatttcgg 240
tgtctgtcat attcgatcaa gtgtgtcata gagcttcctg tttcatctcc cagttattca 300
aggagaggct ggtggctcca cttcccaagg aactgtgctg tgaagatctg aagacaggca 360
cgggctcagg caccgcttgc ctggaatgac aatttggaaac tttaaaagca gcgaccatcc 420
agtcatattt ttcctccat tcc 443

<210> 162
<211> 228
<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(228)

<223> n=A,T,C or G

<400> 162

tcgttatcaa aatgaaagac accaaaccat tactggctc taagctgaca gaaaaggagg 60
aagaaatcggt gactagtgg agtaaattt atgcttnctc aggggaacat gaaaaatgcg 120
gacagtatat tcagaaaggc tattccnagc tcaagatata tnatttgaa ctanaaaata 180
tagcanaatt tgagggcctg acagacttct canatacnn tcaagttt 228

<210> 163

<211> 580

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(580)

<223> n=A,T,C or G

<400> 163

acccaaggct acacatcctt ctgtgaaaca gtctcacgga gactctcaga atcccaagaa 60
ttttttcaa ccttcttttg ttttgattct gaagggaaaca tctgatctgc tctcaatgtt 120
tgttcattct tcaattccaa ggctttatTTTt ggaacagact ttgcattca atggcaggct 180
cgaaggcaga tggcttctcg ggaggctctg ctttggaaagt ttgcntgtcc atcaattcta 240
aggctttagn tggaatagaa actttcatTC tgcaaggagc cttcagaaaa ccattcattat 300
caggagactc ttcttaatttt ccatttattt tatctatttcc tttttgatgc gcaggcttgg 360
gtanacacac atccttctgt gaaacagtct cacagagact ctcagaatcc caagaacttt 420
cttcatagtc cttttgtttg gattctgatg ggagtatctc atctgctctc aatgtttgtt 480
cattctcaa ttccaaggct ttatTTGAA cagacttttgc catttcaatg gcaggctcga 540
aggcagatgg ctctcgggaa ggctctgctt tgaaaagttt 580

210> 164

<211> 140

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(140)

<223> n=A,T,C or G

<400> 164

acttatatct ttgggncttgg ggtttctcaa agttcacgac agacataggc actctcacag 60
tatcaagccc atttaccgnc acttcacacc aataactcgcc ccaccgngng ataggntctg 120
ctggnaactt taatgnatgn 140

<210> 165

<211> 370

<212> DNA

<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(370)
<223> n=A,T,C or G

<400> 165
acatggagcc actgccacca gtggtgatgg aaagcactgc cttcttactc cggaagggtc 60
ctttgtcata catggcagcg taagtgtaaag caaactctcc tatgaacact cgctcaaacc 120
agcctttcag aatggcaggg actccaaacc actgcnnngg ggaactggaa tatcacaagg 180
tctgcggctt ccagttttt ttgttcagcc acaatatctg ggctcanatg gncttctta 240
taagccagaa cagactcggn aggataactga aagttcgcag ggncccttcn tttacctgng 300
atgncccttn tggaaatgtat gggattgaag ntcatggnat aaagggnccga ctncaccacc 360
tccattcttt 370

<210> 166
<211> 258
<212> DNA
<213> Homo sapiens

<400> 166
gtcaaaaagtc atgattttta tcttagttct tcattactgc attgaaaagg aaaacctgtc 60
tgagaaaaatg cctgacagtt taatttaaaa ctatggtta agtctttgac aagaaaaaaaa 120
aacaacacaaa cacttcttcc catcagtaac actggcaatc ttccctgttaa ccactctcct 180
tagggatggt atctgaaaca acaatggtca ccctcttgag attcgttttta agtgttaattc 240
cataatgagc agaggtgt 258

<210> 167
<211> 345
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(345)
<223> n=A,T,C or G

<400> 167
ggtcagccaa acacccagga tctctgtaaa actgaagaac aggncaatgc caccaacaaa 60
tctcaaaacc tctccagcat atttccctat gattggagca catggngagc acnантggtc 120
acttttaaca canctagcca gacaggngnc atttgggtta acacttcggc acccacagca 180
ntttanantt ctctggatgt catttcgagc actttgtatattt attggtcann tttctgtatc 240
tngcgccttgg ttagccctga accaggagca acagggncag cttctggagg ntggttggaa 300
caatacggca agtgnntngaa atgacatccca acctncngaa atgac 345

<210> 168
<211> 61
<212> DNA
<213> Homo sapiens

<400> 168
gatagtgtgg tttatggact gaggtcaaaa tctaagaagt ttgcgcagacc tgacatccag 60
t 61

<210> 169

<211> 344
<212> DNA
<213> Homo sapiens

<400> 169
acattgggc tataaatata aatgctactt atgaagcatg aaattaagct tctttttct 60
tcaagttttt tctcttgc tgcataccaa cgttactcca aacaataaaa aatctatcat ttctgctctg 120
tgctgaggaa tggaaaatga aaccccccacc ccctgaccgg taggactata cagtggaaac 180
tgttcattgc tgatgaatgc agcagtcacc aaaaaataca cccaatctc cagataaccc 240
cagtcactt taggaaatca aaaattaccc ggaagcaatt tagt 300
344

<210> 170
<211> 114
<212> DNA
<213> Homo sapiens

<400> 170
agcagtgtgt cctccatgaa taaacaggag ttctggaggc ccattttctg catttctgc 60
tgattgtct tcccaattt tacttaatc ccacacattc aggcggcggt cagt 114

<210> 171
<211> 150
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(150)
<223> n=A,T,C or G

<400> 171
actgagagca tttataatct gaccaaattc ataggcatta ttaggcttgg ctatcggaag 60
tttctcaggc tcttctggng acctgctgct tttgcctccc ttctcanaag caaggcatcc 120
catggagacc tcccctgcag ggcttccagg 150

<210> 172
<211> 435
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(435)
<223> n=A,T,C or G

<400> 172
atttttttc cactgcctca cactagttag ctgtgccaag tagtagtgg acacccgtgt 60
tgtcatttcc cacatcacgt aagagcttcc aaggaaagcc aaatcccaaga tgagtctcag 120
agagggatca atatgtccat gattatctt tgggttaggt ctacagtcaa tgtgtatgg 180
gtctttgtttt cccagtcgc cagaatatct ttgtgcttct ctaatcatgg gctttaaagc 240
taatcaatgt gttggcagca tctctgtcac tcttggtaa cacgtgaaga aatcaggttag 300
atttttttctt gtggcattgt tttcgaccc aaaatcagggt atgctgacta tttccaagg 360
gtttttcagt tgcttcattt gcttggtaaag cagggaaatcc tcttggntgct tttcttttc 420
tcgatgagcc cgtgt 435

<210> 173

<211> 622
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(622)
<223> n=A,T,C or G

<400> 173
actgnttcc cccaagtcca tgacatgtat acataattaa tggtttgct ccttgattgt 60
tttctccaac atccagacat agaggctgac caacgcttt aatgtatcca gatataacag 120
gattaaggc tggcacatac acctctggat aaatgttggt cagataccat gtaaaaatttt 180
tacactgaag ggggtgtttt atttcaaattc ttttggaaag atcaccaaaat gcttttggtt 240
taacaatttt tgctgcattt gtatttctcc tataaaatat ttcccttgat tcatccatcc 300
agacttctgc aaggcgaact tggttcttag caatcacctg agtgcctttt ggaaagctat 360
gagggctttt gctgcgaaaa acatgtccaa caacagagca aggcataatc tccaactgcc 420
caccacattt ccatactctg aaagacattt ctatatttc acctcccaag atttccattt 480
cttcatcata gcttccaaa tactcaaaat attctttga tatggaaaaa agtccctctg 540
caaaagtggg tggtaattt gggtagggtt catcttcct tcttgcttc tcatgatcag 600
gaagcgactt ccacccaatg aa 622

<210> 174
<211> 362
<212> DNA
<213> Homo sapiens

<400> 174
acgggtgcagt tgacccactg ttggctctcc ttgcagttcc tgatatgtca tcttttagcat 60
gtggctactt acgtaatctt acctggacac tttctaatct ttgcccac aagaatcctg 120
caccccccattt agatgctgtt gagcagatcc ttccttacattt agttcagctc ctgcacatcatg 180
atgatccaga agtggtagca gatacctgct gggcttattt ctaccttact gatggtccaa 240
atgaacaat tggcatgggtt gtggaaaacag gagttgtgcc ccaacttgc aagcttcttag 300
gagcttctga attgccaattt gtgactcctg ccctaaagagc catagggaaat attgtcaactg 360
gt 362

<210> 175
<211> 486
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(486)
<223> n=A,T,C or G

<400> 175
acagnttc tactacactc agcctttat gtgccaagtt tttcttaag caatgagaaa 60
ttgctcatgt tcttcattt ctcaaatcat cagaggccga agaaaaacac ttggctgt 120
tctaaaactt gacacagtca atagaatgaa gaaaattttaga gtatgttattt gattatttca 180
gctcttgcattt tggctccctt ggtgccttctt gagtttgcattt ctcccaaga gagaaccaa 240
tttcttaagat gactggattt cagaagactc ggggacaaca tttgatccaa gatcttaat 300
gttataattttaa accatgtt cagcaatgtt ctattttttt cattttggaa aatctccata 360
atttcaattt gtaaactttt gtaagacctt tctacattttt tttatgttgc tgacttgagt 420
aatgttatca acgtttttgtt aaatatttac tatgtttttc tattagctaa attccaacaa 480
ttttgtt 486

<210> 176
<211> 461
<212> DNA
<213> Homo sapiens

<400> 176
accctggcca ctctttctt tttggctggc caatgtctcc tctgttaggct ccagaaggct 60
ctcaggatg caggcggct cctgcagggt tgagttgcaa tggaaacaaa gacagctgtg 120
gtccccatgc accctcatct ggtgacatcc tgctactgac agtcaaaaga agccttccca 180
gatgaaattt tagtctctg cgcagccatg ctcttcttcc agcaaaagag ccatgtgcag 240
tcgggtctgc tccccatggg ggctttgatg tggcccaagc agtggatctag ccttccagac 300
acgctcaact ctgcacactc ttctggccgc ctcaaggctt ccaggaccct cccgagccct 360
atcagagtcc ttaccctca ggcatactgat accttgcgtgg gtgaccttgg acagattcac 420
ttacctggac teagtttcat aatatgaaaa tgatagggtt g 461

<210> 177
<211> 234
<212> DNA
<213> Homo sapiens

<400> 177
acacattttg taattacctt ttttgtgtt ttgttagcaac catttgtaaa acattccaaa 60
taattccaca gtcctgaagc agcaatcgaa tcccttctc acttttggaa ggtgactttt 120
cacctaattg catattcccc tctccataga ggagaggaaa aggtgttagc ctgccttacc 180
gagagccaaa cagagcccag ggagactccg ctgtggaaa cctcattgtt ctgt 234

<210> 178
<211> 657
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(657)
<223> n=A,T,C or G

<400> 178
gagctcggan ccctagtaac ggccgcccagg gtgctggnat gngcccttgc gagcgngncg 60
ccggggcagg nacttnatc cccccctcatc ttctctgtgc tcattttgtnt ctctcatttt 120
ttggcatatt ttcaagtca cacttaaaaa ctctccatg tattcacttc tcatacttg 180
gtctacatgc cgaacctaag gtcaggattc caaaaagatg agtatecctt caaacgcctc 240
ctaaggctct ggtatacatg actttggctg tgcaacttcat ttagacttca ccttttgg 300
tgctgttgtt ttttacacta gatcccttgc tcttcattaa agataatgaa agattcacat 360
cacagtgcag ctcttcgctt tgcccttcg taagtcggta gcaactgcgg agagttctgg 420
tctgcttaggc atgtgtgaaa tccgctttgtt ggctctctgt gatgggttcc gcttaacgtt 480
tttatttgtc ttatattacac atgccaagggt ggcaacgtga aaaatgtctc tgacgttatt 540
ttccgactgt aaagctgagc attcgatata agtagctgtt ccaatctgtt tgccataact 600
tgccccctgg tcataaggaca ctggcgcttg cctgtgattt gagagctcta ctaatgt 657

<210> 179
<211> 182
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(182)
<223> n=A,T,C or G

<400> 179
acaaaanctt ttaaatttta tattattttgc aaactttgct ttgggtttgt ggcaccctgg 60
ccaccccatc tgctgtgac agcctctgca gtccgtggc tggcagttt ttgatctttt 120
aagtttcctt ccctaccagg tccccatttt ctgtaaggt ttctaggagg tctgttaggt 180
gt 182

<210> 180
<211> 525
<212> DNA
<213> Homo sapiens

<400> 180
acacgtttt ggccccgacc aatgaggcct tcgagaagat ccctagttagt actttgaacc 60
gtatctggc cgacccagaa gcccgtgagag acctgctgaa caaccacatc ttgaagtcag 120
ctatgtgtc tgaagccatc gttcggggc tgctgtaga gaccctggag ggcatgacac 180
tggagggtggg ctgcagcggg gacatgctca ctatcaacgg gaaggcgatc atctccaata 240
aagacatctt acccaccacac ggggtgatcc actacattga tgagctactc atccagact 300
cagccaagac actatggaa ttggctgcag agtctgtatgt gtccacagcc attgaccttt 360
tcagacaagc cggcctcgcc aatcatctct ctgaaagtga gcggttgcacc ctccctggctc 420
ccctgaattt tttttttttt gatggaaaccc ctccaatttga tgccatataca aggaatttgc 480
ttcgaaacca cataattaaa gaccagctgg cctctaagta tctgt 525

<210> 181
<211> 444
<212> DNA
<213> Homo sapiens

<400> 181
acaccacaaat gtgcatcaag gagacgtgcc gattgattcc tgcagtcccg tccatttcca 60
gagatctcg caagccactt accttcccg atggatgcac attgcctgca gggatcaccg 120
tggttcttag tattttgggt cttcaccaca atcctgtgt ctggaaaaac ccaaaggctct 180
ctgacccctt gaggttctct caggagaattt ctgatcagag acaccctat gcctacttac 240
cattctcagc tggatcaagg aactgcattt ggcaggagtt tgccatgattt gaggtaaagg 300
taaccatttgc ctgtattctg ctccacttca gatgtactcc agacccacc acccttta 360
cttcccccaa ccattttatc ctcaagccca agaatggat gtatttgcac ctgaagaaac 420
tctctgaatgtt ttagatctca gggt 444

<210> 182
<211> 441
<212> DNA
<213> Homo sapiens

<400> 182
acaaccttta ttgcttctcc agcattttcc agaagaatgg tgcatttgc gggccacagg 60
ggatgggggaa gtaaaaaata acataaaacga actgaacaga aatgcaggag ggtggcaaga 120
ggggccgaga ttgggtgttc agggcagaga ggtggaaagac cagggggcaat cagtgcttct 180
tagcttcagc ccaccagagt ggagaatttgc tcaacccaa ttttgcgtc cccatctttt 240
tctccagcag ccatcagcat ttgggttctt ttagcagaca ggtctctggc atctggggag 300
aagccttta ggtatcaatcc cagctcatcc tcctcgatga agccacttttgc tccttgc 360
gcatgtaaa cacccatcc acatcatcc cactttttt cttcaggccg accatttgg 420
agaactttt ttgggtcgaa g 441

<210> 183
<211> 339
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(339)
<223> n=A,T,C or G

<400> 183
tgtntcatcn taaggggatt gggctctaga tctgtcgacg gcgcatttag gatttgcnat 60
cggttangt gtcccgaggt catgaatttt tgctctggag cgttattgtt tgtgaagttt 120
atccaggaga gaactatgtat tgtgtcgatg cgtttactgc aggaaganc acggctctcag 180
tcacggaggt gtaagggtgg actgactgan tgagacaagg gatatntngt tnttatannc 240
ttgtgtatgaa cctgcctacc gtttatgtct ctttgcataat gggctctcng tnctgttnatt 300
cncncaagct gcgggggctt ccnccggttctt gggctctga 339

<210> 184
<211> 490
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(490)
<223> n=A,T,C or G

<400> 184
atatacgcaag cttgtacgac cgacacatac ggcgcattgt gctggattgc ttatcttgc 60
gcgcgacgctc tatataancg anactacata gtctcgaaaa tccactcant ttcaagttcc 120
caaaaanacng ganaaaaaacc catgccttat ttaactaanc atcagctcgc ttctccttct 180
gtaaccgcgc ttntngctcc cagcctatag aagggtaaaa cccacactcg tgcgnccagtc 240
atcnnataac tgattcggccg gggtaactgccc gggggccgct cganaccaat tngcanaatt 300
cacacattgc ggcgctcnan aagctctaga agggcaatcg ccatattgtat ctatacatta 360
tggccgtcgtnacacgtcg tgacgggana ncctggngta ccattaatcg ctgcacanc 420
ccttcgcagc tggggtnac aaaagccgccc catcnctcca cgttgcgncc gatggcaagg 480
acnccctnat 490

<210> 185
<211> 368
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(368)
<223> n=A,T,C or G

<400> 185
ctnnanatag cangcttgcgta cgaccgacac aatacggcca ntgtgctgga ttcgcattcag 60
cgccgccccgg gcagtaccgg cgctcatcta tcngatgatg ggcaccaat gtggggtttt 120
aacctttta tatggctggg gacanaaaagc gcggtacnn aaccnataac gagctgatgg 180
tcatttaaaa atgcttgggg ttttcccggt ctttgggaa attgaaactg agtggactt 240

canaaaactgt gctactttcg cttatctaag tactcgccg caacacccat ccgaatccgc 300
anatatcatc acnctggcg gcgtcancat gcntctaaag ggccaattcn cctanatgag 360
tcttatac 368

<210> 186
<211> 214
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(214)
<223> n=A,T,C or G

<400> 186
ngggagatcg cagcttgtac gactcgcat ataacgnnca atgtgctgga tcgccttcanc 60
gcccggcg gctaatctg gttcgattn tgggtgtntt gtctntnta canggtgcta 120
tcccccttctt cctccttcctc tgccatcctc atcctttatc tccttttgg acaagtgtca 180
nancagacag angcagggtg gtggcaccgt tgaa 214

<210> 187
<211> 630
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(630)
<223> n=A,T,C or G

<400> 187
cagctggac gagtcgatca tatacgccgc atgtgttgn a tcgcatacg gtccggcgag 60
tanttattan attactgtta tttctgtcc tactggatat gatctcttga nggcangtct 120
gtgtcgctg gtcacaccat gttctcaggc tggcaataa ctttcctata atagttatg 180
gataatgaat gacgactang tctanaaaana cgctagctaa ataacacact caggaaaga 240
gtcttaataa ttgtgaaggt gtttttanta tacaacnttt gtttacataa taggaaataa 300
tttttagact tttaaacaga cacttgagcc agattttgtt atgttaccat ctatagtgtc 360
ttgaaaatat tcctcttagt ttccaatatg aatgaatcta aaatccatct ttcaattat 420
gcccaggccc gtgggtcaatg cnccctcnac acttcattaa cggattatac cttggaaac 480
cataatctgg cntagggacga atcgccctggc ncangctaann aactgcccctg tattgagggg 540
ttatnctga ttgcngaggt gctctccag gtccccaaag ggtcgtaactg ttgaanctgg 600
ctctaatttt ntcttgctn acaggtctcc 630

<210> 188
<211> 441
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(441)
<223> n=A,T,C or G

<400> 188
cnngcaanac anggtcggt tccgntgagg naanaattcc ctnataggc tcgcccccta 60

ttcaccaaac caancngaaa ctcttgcgtt caaatctaag ctatnnacaca accccactct 120
 gnaggtatg cgccccgccc ctgcaatgaa atcaatanca tatttggaga cagagagata 180
 gagagagaga ggttcctggc cttnnctatt ctgtcttac ttgnnagatn tcaganatag 240
 aaaaacctat cctaggtccn nccaatgatn gcggcttncg aatcccnng tggccantcc 300
 ccggatcga ctaaatcaaa gaagatcctc cgtcncctg ttccctccaca ctggagtccc 360
 attgtatgc tgggtntttc actggctnat cataccnnag gatctgtcca ccttnaactc 420
 ttctctngga antccctncc c 441

<210> 189
 <211> 637
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(637)
 <223> n=A,T,C or G

<400> 189
 agggngtata tacccacttg tacnactcga tcatanacgc gcatntctga atcgcttnc 60
 ggccgcgtg tactgtggc acttaagcac tgagtactgt ttgcgtcatg ccngtcan 120
 agatgtctg ccaaaggac tccaaacnnaaa tacactgtct tcaacaggag ttaacacctc 180
 acactggtg ganaanagaa ctcactggggt gtgatgcaca cgactgnatc catcaagtgc 240
 gtttgcctgt tgactgctaa ccaaggctct ggcaatgtt gcccggccg cgctcgaaac 300
 caaatctgca aatatcatca cactggccgn cgctcagcat catctanaag gccatcgcc 360
 atagttagtc tatacatcat ggccgcntt acactcctac tggaaaacct gcttaccact 420
 taatgcgttc acacatcccc ttgcngtn gcttatanca aaaagccac gatgcctcca 480
 cattgcncnc tgatggcatg ancccttac gcatancc ggggtntgt tacncangt 540
 accgtctgc acgctacncc ttttccttct ccttcccccc ttcccggttcc tcaccattcg 600
 gggcctttagg tcnatatctc gnccacccaa atntagg 637

<210> 190
 <211> 653
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(653)
 <223> n=A,T,C or G

<400> 190
 agggggtata tacccacttg tacgactgna tcatatacgc gcatgtctgg aatcgcttnc 60
 gtggctcca ttttgcata ctacttctaa gaactacaaa agtgatactg angatacatt 120
 acacagaang gctnacattc tcncagatcc tcattntca tgatatgtgg acatcangan 180
 cacgtggata agtgtatcta aanaatggct ttccacttta ttaaggttt 240
 acatganatt cttttttttt cttttttttt taacatctaa tcggaaaacta 300
 tgcctnaact gcaacntttt tttttttttt taacatctaa tcggaaaacta 360
 ccnaatctgc gatcctcac ctggcgccgc tcaacatcat ctaaaggcca atcgccctata 420
 ntantctata cttttttttt gcttaatggg aaacccgggtt accacttata 480
 gtttgcgttca cttttttttt tttttttttt taacatctaa tcggaaaacta 540
 canctgatgc aatgacccct gttttttttt tttttttttt taacatctaa tcggaaaacta 600
 cagcgtgcn cttttttttt gctttttttt tttttttttt taacatctaa tcggaaaacta 653

<210> 191

<211> 663
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(663)
<223> n=A,T,C or G

<400> 191
anggnntata tacccactgt ncgactcgat catatacgcg catgtcgat cggctccanc 60
gcgcggcat gtactatata tacatcaact gtattatcat ttanatattg atnaaagaca 120
aaatcatact tccatctgct cactgatgt aattactatg atacatgatc atgtaaacgt 180
atcaatataa caatggaaga tccctctgac tatgcaagcc taatttcca atcnccatgca 240
ctctcatagc tcaaananatnt cacngacatc ctgatgaaac tatnatacan tttccacaca 300
aatcacttcg ctttagatct ctccattatt ctgctttc cccccctaaca actacaatc 360
ctcnntggat gggagaata tatatcatct actaaaaata atatataatc ccctgcanat 420
ttgtggnaaa tcnggtgtct caanagccac aggagnacaa ggggnacca actaggactt 480
ttgtatgctt atctctgtac tcgcgcacac ctaagcgatt ctgcnattct ccctggcgcc 540
gtcacanctc tanaggccat cncnatatga tctatacatc ntggcgtctt tacactctga 600
cgaaaaacgg gtnccantta ccctggacca tccctcgcn ctgntataca aagccccga 660
ncc 663

<210> 192
<211> 361
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(361)
<223> n=A,T,C or G

<400> 192
anttttata tacccactgg tacaactcgaa ncctatacgg cgcanttncg gaatcanctt 60
cancggcgcc ggcattgtacc ggtnatcatc atcngatgtat ggcgctcnaa tgtgggtttt 120
acctnttata cggctgagat canatcgctt acataacaaa nncaactgtat ggttaatnta 180
aatnccgttg gtttctcccn ntctgttggg gaacttgana ctgagtgngat cttccatana 240
cgtgctattt tcggctancn antcctcagc gnacacctat ngnagtgcgc naattcatcc 300
atgntggctt cgactnttcc aaaangccnt ncggccacnt gntcgcnana cantctcggc 360
c 361

<210> 193
<211> 314
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(314)
<223> n=A,T,C or G

<400> 193
aggggnntata taccaactgg tncgactcgaa tcctatacgc gcatttcgga ttgcgttcaa 60
cggcgccggc atgtacaaa cctcaatccc aaccgtctca ntngacggg ctcagttctg 120
tcacagccac cccacatttc tttgttttgc tctgccactt caaaagaatt ccaaataaga 180

attctgctgc agctccgtac aaggatatgg gcagcacagc acacacagag tngtgctcct 240
 cacacttctc tggnaatgtc tcgtaatat ctcaacagtc angaagtggg gcgttatcaa 300
 aaacaatcag ggcc 314

<210> 194
 <211> 550
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(550)
 <223> n=A,T,C or G

<400> 194
 aggnngnata taccactgg tncgactcga tcctatacgc gcatgtcgg ncgttatgtg 60
 gtcncgcaag taccttctc gcagtatgg tctgtntcct ctatgatnag tgatcgaata 120
 atcatcgaaat tcancgaaag ttattcgagt gatatntgtg gctttagaa tctatgctcc 180
 atgggtgggt cactgtcaag attaacacag aatggaaagan ncngcactgc ataaaagatg 240
 ttgtcaaatt gggtgcgtt atcngatagc tcntcccaag aggtcantgg tgttcaggat 300
 tncnacataa gatnttggat caccnacga ccagangata ccngtgc当地 ctgtgaancn 360
 ngtatctgc ctatncctgc cctctcgan gatcccctcg ggacgacgag atcattctgg 420
 aaacagcnan ttagatgtcca gtnnangatt gatgancgac ganacgc当地 atanatgtct 480
 gacgtgagat tnggatgtga atcttccnt gtgtgacctg cnccntacnn aanggtgcgn 540
 ctccactcnn 550

<210> 195
 <211> 452
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(452)
 <223> n=A,T,C or G

<400> 195
 nngcgggnat gataccaact ggtacgaact cganctctat nacggcgtt tttcnngatc 60
 tgctatgtgg tctcggaat gtacattata acnnggcana catataatct acnctgtct 120
 ttntctcccc cngagagcgc aancatctcc aaatcgggtt ctgggtcata caatggtctc 180
 cantaatcac acaactcata tatattttagt gaangtgtct gtcatcgatcc ccacgangga 240
 agtnncgtcg ctgtntgtct gtcacttaggt gngtactctc cagtaactga aancgtgtta 300
 nngctgtcg tngtactggc cggcgc当地 cgaat ctgttnatcatcacaatng 360
 cgnccccc当地 ncatcaactna gggncanttc gcctataactg atcgtnlgcg annccctgcgn 420
 cncttacacg tcgnacgggaa naccggcctt cc 452

<210> 196
 <211> 429
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(429)
 <223> n=A,T,C or G

<400> 196

gcgggnnat gataccagct ngtacgactc gatcctataa cggcgcatgt gngtatccgc 60
 tacgtgtctc ggcgtatgtac atataacggg gcaacatata atnatacant ctgtcttttt 120
 ctccccggaa aacggcaacc atctccaata tcggtctggg tctccaatgg tctccaacta 180
 aatcacacaa gtcaaatata ntangaaa gtgtctgtct cttcccccaga aggagtancg 240
 ttagctgttg tctgtcatta gggtggtacc tccagtnaca tgaaaactgg tgagggtgtc 300
 cttgtacaag ctctgcctca ccagatccta tactattagg gggcccacgg ttatctatct 360
 taagggtctn aaaacctgga cttcatctgc tccggcggan gaatgtcccg cttaacttacg 420
 ntgttccac 429

<210> 197

<211> 471
 <212> DNA
 <213> Homo sapiens

<220>

<221> misc_feature
 <222> (1)...(471)
 <223> n=A,T,C or G

<400> 197

atgataacgca gctngtacga gccgtcacta tnacggcnca ttgtgtggat tcngctntga 60
 tcggcgcccc ggcgtatgtcca tcnagagcgc atcatgggan tgnactcccc atatnntgac 120
 caangttcgc gcaaggagcc naganccgat actacctgag ctgtcgtctn gttatacacg 180
 tttctggcca angancaact ccacatncaa caagttggtg ttgaaatgtt gtttatnagt 240
 ccaccaacccg gccgctctgt cccttcccgta tgatccgaag ataagcttcc tgtccggaan 300
 acgaacggcg tgggtgtnggg acatantgat atgtcggggt caggaagtac tcgnncgcaac 360
 ncgcaagcna atctgcnata tcatcacctg gccgcgtcg agctgccana ngccnttcg 420
 cctatatgag tctatacatt cctggccgtc nttaacactc ngacggggaaa c 471

<210> 198

<211> 643
 <212> DNA
 <213> Homo sapiens

<220>

<221> misc_feature
 <222> (1)...(643)
 <223> n=A,T,C or G

<400> 198

tngtncgacc gtcactatac gcccattgtgtt ggtatccgntc cacggcgccg ggcangtacg 60
 anactatatt gatcctctga tattgaaagt tggctanca ataaccttta angcaaatac 120
 ctcantgagt tttgaccaga agtcaccaca tcatgaatca cagtctatgg caaatgatac 180
 cagtgtctct aagtcttatac ctcaggtaa gagcatgcta ttccgtttta catttactgg 240
 aattttactgt tcattcatna ttaaaaatctc tagtttcat cctcaactgt ctaanaccag 300
 tggcacaga cttaaagactc tggctctctc attttctcca acagaaacat tctcagtgtc 360
 tactgttctta aaaggaaatt tccgagggtgg cacttctcg aatatacgacc ctcnggctct 420
 atcagggtt acttcnngca ctgcgtcattt ggggttggc anttgcattt tctgtccagt 480
 cacttcattt taagaaaaca attgatcgct ggtcacatgt nattcattgg cagccgggt 540
 gactgtcgag tctcgcgcac acnctagcaa tcgnnattct ccatggngcg tcactctcta 600
 naggccatcc cctatatgat ctataatctg gctgtttac act 643

<210> 199

<211> 292
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(292)
<223> n=A,T,C or G

<400> 199
ncggcnggag ttcgcagttt nacgaccgat cctatacgnc gcatttctga tccgctacnt 60
gtccggcgag tctatgttat ttattnntga taaaatcaat attttcttgc tgaatattaa 120
tcttatctnt acttttatac tattgaccta gctatatgtt ttgancttt tgaactccta 180
tcagtnnttt tcatgtatc gtatattttc cacttggtac ctntngctga ntccttagata 240
tcgtaaaaaca tctctnnatc ntcacacnngt gnccaggnt ctgtatngaa tt 292

<210> 200
<211> 275
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(275)
<223> n=A,T,C or G

<400> 200
atacgcaagc ttggtaccga gctnggatcc ctattaaccg gccgcaatat tctggaattc 60
tgcttancgt ggtcncggcc gaagtactat gctatnttac tttttggga tataaaatca 120
atataatttct ttctnaagta tataaatctt atccncgtat ctttcnatac ctntctgaca 180
ntaagcttat angtatntga tctntgttga actcctatca agtgnttctn catgctatcg 240
tganntcttc cacnttggta cttttacgc tgaat 275

<210> 201
<211> 284
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(284)
<223> n=A,T,C or G

<400> 201
cgnnnatcca gtgtanaccg tcnttacgcg cattctgate gttcacgccc gcgtctttat 60
atctatctcg actgattcac ctgtcattgt aaanaattcg tgcagctgt ctaccnctta 120
nacatcatct aatcnaacta ncctgataaa tttcttcaat agggatanaac ntntagtaca 180
tacgnnttcca ttgagntacn tccgcggacc cncatcgaa acnnncatgcg gtcagtcnna 240
gcatcctcta tcttaatccg tccttacnt ntgaacgctc cact 284

<210> 202
<211> 448
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<222> (1)...(448)
<223> n=A,T,C or G

<400> 202
atgatacgca agcttgcgtacg actcggtatca tataacggcc gcaatgtgtc ggaattccgc 60
ttcgacggac gccgggcgtg tactttata atnctactcc tcagacccctg catctcnacc 120
gctnggtcca gtttgtaaaa acnnacttcc gtngtgcgc cctggttctg ancantctct 180
atcacnctct atcctcnccat ccncaanact anatcgcggtg aattcatatt tattcatttt 240
ccataatgtat gggggaaangaa ctatcnctna tnatgtttt cacnctngtgc gcanntcgnc 300
natctcgcnna ncgtgaaac gattactctg tcgcgaaccc tctangntga attctgcnaa 360
atatctntna cnctggcngg cgctcnangn atgcctctcg anggccaatc cgccnngcat 420
gattctaaatt anatccnntng gtcccnntt 448

<210> 203
<211> 321
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(321)
<223> n=A,T,C or G

<400> 203
gggtgcnaga tcgcagtngt acgaatcgnt catatacgcc gcatgtgntg antcgctacg 60
tgtccggcga ngtaccatataatcgaanta ncatagttctt ggangccnc tcattttcaa 120
tttccccaaaa nacgggaaaaa ccnaagcctt attaactaa ctatctgctc gcttctcgct 180
tctgtaccgc gctatntgct nccagcctt aanaaggta aaacccacac tcgggtcgctc 240
agtctccnat atantgagtc nccgggtact ggccggcgg tcgttcnaaa ncaattcncc 300
aanttcaacta ctggcgccgc c 321

<210> 204
<211> 369
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(369)
<223> n=A,T,C or G

<400> 204
ntgtngttag tacccagtgg tacgactcgatccatgtacg ggcgcgtgtc ctgaatcggtt 60
acttgcgtcg gccaagtatc tataaagccaa actatcacag ttctgaaaatg ccatatcattt 120
ttcagttccc aaaagancgg gaaaacccaa gccttattaa actaacaatc agtgcgttc 180
gcttcgtac cgcgttttgc gccccccagcc tataaaaggta taaaacccac actcggtcg 240
ccagtcattcg ataactgaat cgcgggtac tgcccggcgc ggcgtcnann ccaaattctgc 300
agatattcaca cactggcgcc gtcancatcg ctctagaagg ccaattcncc tatantgatt 360
ctattacaa 369

<210> 205
<211> 2996
<212> DNA
<213> Homo sapien

<400> 205

cagccaccgg	agtggatgcc	atctgcaccc	accgcccgtga	ccccacaggc	cctgggctgg	60
acagagagca	gctgtatttg	gagctgagcc	agctgaccca	cagcatca	gagctgggccc	120
cctacaccct	ggacaggac	agtctctatg	tcaatgggtt	cacacagcgg	agctctgtgc	180
ccaccactag	cattcctggg	accccccacag	tggacctggg	aacatctggg	actccagttt	240
ctaaacctgg	tccctcggt	gccagccctc	tccctgggtct	attcaactctc	aacttcacca	300
tcaccaacct	geggatgag	gagaacatgc	agcacccctgg	ctccagaag	ttcaacacca	360
cggagagggt	cttcagggc	ctggccctgt	ttcaagagca	ccagtgttgg	ccctctgtac	420
tctggctgca	gactgactt	gctcaggct	gaaaaggatg	ggacaggcac	tggagtggat	480
gccatctgca	cccaccaccc	tgacccaaa	agccctaggc	tggacagaga	gcagctgtat	540
tggagactga	gccagctgac	ccacaatata	actgagctgg	gccctatgc	cctggacaac	600
gacagectct	ttgtcaatgg	tttcaactat	cgagactctg	tgtccaccac	cagcaactct	660
gggacccccc	cagtatct	gggagcatct	aagactccag	cctcgatatt	tggcccttca	720
gctccagcc	atctctgtat	actattcacc	ctcaacttca	ccatcaactaa	cctcggttat	780
gaggagaaca	tgtggcctgg	ctccaggaag	ttcaacacta	cagagagggt	ccttcagggc	840
ctgctaaggc	ctttgttcaa	gaacaccagg	gttggccctc	tgtactctgg	ctgcaggctg	900
acttgctca	ggccagagaa	agatggggaa	gocacccggag	tggatccat	ctgcacccac	960
cgcctgacc	ccacaggccc	tgggtggac	agagagcagc	tgtatttgg	gctgagccag	1020
ctgacccaca	gcatca	gctggcccc	tacacactgg	acagggacag	tctctatgtc	1080
aatgtttca	cccatcgag	ctctgtaccc	accaccagg	ccgggttgg	cagcgaggag	1140
ccattcacac	tgaacttca	catcaacaac	ctgcgttaca	tggcggacat	ggcccaaccc	1200
ggctccctca	agttcaacat	cacagacaac	gtcatgaagc	acctgctcag	tcctttgttc	1260
cagaggagca	gcctgggtgc	acggtacaca	ggtgcaggg	tcatcgact	aagtctgtg	1320
aagaacggtg	ctgagacacg	ggtggacctc	ctctgcacct	acctgcagcc	cctcagcggc	1380
ccaggtctgc	ctatcaagca	ggtgttccat	gagctgagcc	agcagaccca	tggcatcacc	1440
cggctggcc	cctactctt	ggacaaagac	agccttacc	ttaacggtt	caatgaaccc	1500
ggtccagatg	agccttctac	aactccaa	ccagccacca	cattccgtcc	tcctctgtca	1560
gaagccacaa	cagccatggg	gtaccaccc	aagaccctca	cactcaactt	caccatctcc	1620
aatctccagt	attcaccaga	tatggcaag	ggctcagct	cattcaactc	caccgagggg	1680
gtccttcage	acctgctcag	acccttgc	cagaagagca	gcatggccc	cttctacttg	1740
ggttccaa	tgatctccct	caggcctgag	aaggatgggg	cagccactgg	tgtggacacc	1800
acctgcaccc	accaccctga	ccctgtggc	cccggtctgg	acatacagca	gcttactgg	1860
gagctgagtc	agctgaccca	tgggtcacc	caactgggt	tctatgtct	ggacagggat	1920
agcctttca	tcaatggct	tgcacccag	aatttatcaa	tccggggcga	gtaccagata	1980
aatttccaca	ttgtcaactg	gaacctcagt	aatccagacc	ccacatcc	agagtacatc	2040
accctgctga	gggacatcca	ggacaaggc	accacactct	acaaaggcag	tcaactacat	2100
gacacattcc	gcttctgcct	ggtcaccaac	ttgacgatgg	actccgtgtt	ggtactgtc	2160
aaggcattgt	tctcttccaa	tttggacccc	agcctgggt	agcaagtctt	tctagataag	2220
accctgaatg	cctcattcca	ttggctggc	tccacctacc	agttgggt	catccatgt	2280
acaaaaatgg	agtcatcagt	ttatcaacca	acaaggcag	ccagcaccc	gcacttctac	2340
ctgaatttca	ccatcaccaa	cctaccat	tccaggaca	aagccagcc	aggaccacc	2400
aattaccaga	gaaacaaaag	gaatatttg	gatgcgtca	accaacttt	ccgaaaacagc	2460
agcatcaaga	gttatttttc	tgactgtca	gttcaacat	tcaggtctgt	ccccaacagg	2520
caccacccg	gggtggactc	cctgtgttac	ttctcgcc	tggctggag	agtagacaga	2580
gttgcacat	atgaggat	tctgaggat	accggaaatg	gtacccagct	gcagaacttc	2640
accctggaca	ggagcgtgt	ccttggat	gggttattttc	ccaacagaaa	ttagccctt	2700
actggaaat	ctgaccccttcc	cttctgggt	gtcatectca	tgggttggc	aggactctg	2760
ggactcatca	catgcctgt	ctgcgggt	ctgggtacca	ccggccggcg	gaagaaggaa	2820
ggagaataca	acgtccagca	acagtgc	ggctactacc	agtcacac	agacctggag	2880
gatctgcaat	gactggaaact	tgcgggt	tgggtgcct	ttccccccagc	cagggtccaa	2940
agaagcttgg	ctggggcaga	aataaaccat	attggtcgga	cacaaaaaaa	aaaaaaa	2996

<210> 206

<211> 914

<212> PRT

<213> Homo sapien

<400> 206

Met Ser Met Val Ser His Ser Gly Ala Leu Cys Pro Pro Leu Ala Phe
 1 5 10 15
 Leu Gly Pro Pro Gln Trp Thr Trp Glu His Leu Gly Leu Gln Phe Leu
 20 25 30
 Asn Leu Val Pro Arg Leu Pro Ala Leu Ser Trp Cys Tyr Ser Leu Ser
 35 40 45
 Thr Ser Pro Ser Pro Thr Cys Gly Met Arg Arg Thr Cys Ser Thr Leu
 50 55 60
 Ala Pro Gly Ser Ser Thr Pro Arg Arg Gly Ser Phe Arg Ala Trp Ser
 65 70 75 80
 Leu Phe Lys Ser Thr Ser Val Gly Pro Leu Tyr Ser Gly Cys Arg Leu
 85 90 95
 Thr Leu Leu Arg Pro Glu Lys Asp Gly Thr Ala Thr Gly Val Asp Ala
 100 105 110
 Ile Cys Thr His His Pro Asp Pro Lys Ser Pro Arg Leu Asp Arg Glu
 115 120 125
 Gln Leu Tyr Trp Glu Leu Ser Gln Leu Thr His Asn Ile Thr Glu Leu
 130 135 140
 Gly Pro Tyr Ala Leu Asp Asn Asp Ser Leu Phe Val Asn Gly Phe Thr
 145 150 155 160
 His Arg Ser Ser Val Ser Thr Thr Ser Thr Pro Gly Thr Pro Thr Val
 165 170 175
 Tyr Leu Gly Ala Ser Lys Thr Pro Ala Ser Ile Phe Gly Pro Ser Ala
 180 185 190
 Ala Ser His Leu Leu Ile Leu Phe Thr Leu Asn Phe Thr Ile Thr Asn
 195 200 205
 Leu Arg Tyr Glu Glu Asn Met Trp Pro Gly Ser Arg Lys Phe Asn Thr
 210 215 220
 Thr Glu Arg Val Leu Gln Gly Leu Leu Arg Pro Leu Phe Lys Asn Thr
 225 230 235 240
 Ser Val Gly Pro Leu Tyr Ser Gly Cys Arg Leu Thr Leu Leu Arg Pro
 245 250 255
 Glu Lys Asp Gly Glu Ala Thr Gly Val Asp Ala Ile Cys Thr His Arg
 260 265 270
 Pro Asp Pro Thr Gly Pro Gly Leu Asp Arg Glu Gln Leu Tyr Leu Glu
 275 280 285
 Leu Ser Gln Leu Thr His Ser Ile Thr Glu Leu Gly Pro Tyr Thr Leu
 290 295 300
 Asp Arg Asp Ser Leu Tyr Val Asn Gly Phe Thr His Arg Ser Ser Val
 305 310 315 320
 Pro Thr Thr Ser Thr Gly Val Val Ser Glu Glu Pro Phe Thr Leu Asn
 325 330 335
 Phe Thr Ile Asn Asn Leu Arg Tyr Met Ala Asp Met Gly Gln Pro Gly
 340 345 350
 Ser Leu Lys Phe Asn Ile Thr Asp Asn Val Met Lys His Leu Leu Ser
 355 360 365
 Pro Leu Phe Gln Arg Ser Ser Leu Gly Ala Arg Tyr Thr Gly Cys Arg
 370 375 380
 Val Ile Ala Leu Arg Ser Val Lys Asn Gly Ala Glu Thr Arg Val Asp
 385 390 395 400
 Leu Leu Cys Thr Tyr Leu Gln Pro Leu Ser Gly Pro Gly Leu Pro Ile
 405 410 415
 Lys Gln Val Phe His Glu Leu Ser Gln Gln Thr His Gly Ile Thr Arg
 420 425 430
 Leu Gly Pro Tyr Ser Leu Asp Lys Asp Ser Leu Tyr Leu Asn Gly Tyr
 435 440 445
 Asn Glu Pro Gly Pro Asp Glu Pro Pro Thr Thr Pro Lys Pro Ala Thr

450	455	460
Thr Phe Leu Pro Pro Leu Ser Glu Ala Thr Thr Ala Met Gly Tyr His		
465	470	475
Leu Lys Thr Leu Thr Leu Asn Phe Thr Ile Ser Asn Leu Gln Tyr Ser		480
485	490	495
Pro Asp Met Gly Lys Gly Ser Ala Thr Phe Asn Ser Thr Glu Gly Val		
500	505	510
Leu Gln His Leu Leu Arg Pro Leu Phe Gln Lys Ser Ser Met Gly Pro		
515	520	525
Phe Tyr Leu Gly Cys Gln Leu Ile Ser Leu Arg Pro Glu Lys Asp Gly		
530	535	540
Ala Ala Thr Gly Val Asp Thr Thr Cys Thr Tyr His Pro Asp Pro Val		
545	550	555
Gly Pro Gly Leu Asp Ile Gln Gln Leu Tyr Trp Glu Leu Ser Gln Leu		560
565	570	575
Thr His Gly Val Thr Gln Leu Gly Phe Tyr Val Leu Asp Arg Asp Ser		
580	585	590
Leu Phe Ile Asn Gly Tyr Ala Pro Gln Asn Leu Ser Ile Arg Gly Glu		
595	600	605
Tyr Gln Ile Asn Phe His Ile Val Asn Trp Asn Leu Ser Asn Pro Asp		
610	615	620
Pro Thr Ser Ser Glu Tyr Ile Thr Leu Leu Arg Asp Ile Gln Asp Lys		
625	630	635
Val Thr Thr Leu Tyr Lys Gly Ser Gln Leu His Asp Thr Phe Arg Phe		640
645	650	655
Cys Leu Val Thr Asn Leu Thr Met Asp Ser Val Leu Val Thr Val Lys		
660	665	670
Ala Leu Phe Ser Ser Asn Leu Asp Pro Ser Leu Val Glu Gln Val Phe		
675	680	685
Leu Asp Lys Thr Leu Asn Ala Ser Phe His Trp Leu Gly Ser Thr Tyr		
690	695	700
Gln Leu Val Asp Ile His Val Thr Glu Met Glu Ser Ser Val Tyr Gln		
705	710	715
Pro Thr Ser Ser Ser Thr Gln His Phe Tyr Leu Asn Phe Thr Ile		720
725	730	735
Thr Asn Leu Pro Tyr Ser Gln Asp Lys Ala Gln Pro Gly Thr Thr Asn		
740	745	750
Tyr Gln Arg Asn Lys Arg Asn Ile Glu Asp Ala Leu Asn Gln Leu Phe		
755	760	765
Arg Asn Ser Ser Ile Lys Ser Tyr Phe Ser Asp Cys Gln Val Ser Thr		
770	775	780
Phe Arg Ser Val Pro Asn Arg His His Thr Gly Val Asp Ser Leu Cys		
785	790	795
Asn Phe Ser Pro Leu Ala Arg Arg Val Asp Arg Val Ala Ile Tyr Glu		800
805	810	815
Glu Phe Leu Arg Met Thr Arg Asn Gly Thr Gln Leu Gln Asn Phe Thr		
820	825	830
Leu Asp Arg Ser Ser Val Leu Val Asp Gly Tyr Phe Pro Asn Arg Asn		
835	840	845
Glu Pro Leu Thr Gly Asn Ser Asp Leu Pro Phe Trp Ala Val Ile Leu		
850	855	860
Ile Gly Leu Ala Gly Leu Leu Gly Leu Ile Thr Cys Leu Ile Cys Gly		
865	870	875
Val Leu Val Thr Thr Arg Arg Lys Lys Glu Gly Glu Tyr Asn Val		880
885	890	895
Gln Gln Gln Cys Pro Gly Tyr Tyr Gln Ser His Leu Asp Leu Glu Asp		
900	905	910
Leu Gln		

<210> 207
<211> 2627
<212> DNA
<213> *Homo sapiens*

<400> 207
ccacgcgtcc gcccacgegt ccggaaggca gcggcagctc cactcagcca gtaccaggat 60
acgctggaa cttccccag ccatggctc cctggggcag atcctttct ggagcataat 120
tagcatcatc attattctgg ctggagcaat tgcactcatc attggctttg gtatttcagg 180
gagacactcc atcacagtca ctactgtcgc ctcagctggg aacattgggg aggatggaat 240
cctgagctgc actttgaac ctgacatcaa actttctgtat atcgtgatac aatggctgaa 300
ggaagggttt ttaggcttgg tccatgagtt caaagaaggc aaagatgagc tgcggagca 360
ggatgaaatg ttcaaggccc ggacagcagt gtttgcgtat caagtgtata gttggcaatgc 420
ctcttgcgg ctgaaaaacg tgcaactcac agatgctggc acctacaaat gttatatcat 480
caactctaaa ggcaagggga atgctaacct tgagtataaa actggagcc tcagcatgcc 540
ggaagtgaat gtggactata atgccagctc agagacctt cgggtgtgagg ctccccatg 600
gttccccagg cccacagtgg tctgggcatc ccaagttgac cagggagcca acttctcgga 660
agtctccaaat accagctttg agctgaactc tgagaatgtg accatgaagg ttgtgtctgt 720
gctctacaat gttacgatca acaacacata ctccctgtatg attgaaaatg acattgccaa 780
agcaacaggg gatactaaag tgacagaatc ggagatcaaa aggccggagtc acctacagct 840
gctaaactca aaggttctc tttgtgtctc ttctttttt gccatcagct gggcacttct 900
gcctctcagc ctttacctga tgctaaaata atgtgccttg gcccacaaaaa agcatgcaaa 960
gtcattgtta caacagggat ctacagaact atttcaccac cagatatgac ctatgtttat 1020
atttctggaa ggaatgaat tcatatctag aagtctggag tgagcaaaaca agagcaagaa 1080
acaaaaagaa gccaaaagca gaaggctcca atatgaacaa gataaatcta tcttcaaga 1140
catattagaa gttggaaaaa taattcatgt gaactagaca agtgtgttaa gagtgataag 1200
taaaatgcac gtggagacaa gtgcattcccc agatctcagg gacccccc tgcctgtcac 1260
ctggggagtg agaggacagg atagtgcatg ttctttgtct ctgaatttt agttatatgt 1320
gctgtatgt tgctctgagg aagccctgg aaagtcatac ccaacatatac cacatctt 1380
attccacaaa ttaagctgta gtatgttaccc taagacgctg ctaattgact gccacttcgc 1440
aactcagggg cggctgcatt ttagtaatgg gtcaaatgat tcactttta tgatgtttcc 1500
aaaggtgcct tggcttctct tcccaactga caaatgccaa agttgagaaa aatgatcata 1560
atttagcat aaacagagca gtccggcaca cccgattttat aaataaaactg agcacccct 1620
ttttaaacaa acaaatgccc gtttattttc cagatgtatg tcatccgtga atggtcagg 1680
gaaggaccc tccacccgtac tatatggcat tatgtcatca caagctctga ggcttccct 1740
ttccatcctg cgtggacagc taagacccca gtttcaata gcatcttagag cagtgggact 1800
cagctgggtt gattcgtccc cccatctccg gggaaatgtc tgaaagacaat ttgggttacc 1860
tcaatgaggg agtggaggag gatacagtg tactaccaac tagtggataa aggccaggaa 1920
tgctgctcaa cctcttacca tgcacggac gtctccccat tacaactacc caatccgaag 1980
tgtcaactgt gtcaaggact agaaaccctg gttttgagta gaaaaggccc tggaaagagg 2040
ggagccaaca aatctgtctg cttccatcata tttagtattt gcaaaataacg attctgtctc 2100
tttggctgt gcctcagcac agagagccag aactctatcg ggcaccaggaa taacatctt 2160
cagtgaacag agttgacaag gcctatggga aatgcctgtat gggattatct tcagctgtt 2220
gagttctaa gtttcttcc cttcattctt ccctgcaagc caagttctgt aagagaaatg 2280
cctgagttct agctcagggtt ttcttactctt gatatttagat ctccagaccc ttcttgccca 2340
caattcaaat taaggcaaca aacatataacc ttccatgaaag cacacacaga cttttgaaag 2400
caaggacaat gactgcttga attgaggcct tgaggaatga agctttgaag gaaaagaata 2460
cttggttcc agcccccttc ccacactctt catgtgttta ccactgcctt cctggaccc 2520
ggagccacgg tgactgtatt acatgttgg atagaaaact gatatttagag ttctgtatgt 2580
tcaagaaatgatttataat acatttccca caccaaaaaa aaaaaaaaaa 2627

<210> 208
<211> 282
<212> PRT
<213> *Homo sapiens*

<400> 208

Met Ala Ser Leu Gly Gln Ile Leu Phe Trp Ser Ile Ile Ser Ile Ile
 5 10 15

Ile Ile Leu Ala Gly Ala Ile Ala Leu Ile Ile Gly Phe Gly Ile Ser
 20 25 30

Gly Arg His Ser Ile Thr Val Thr Thr Val Ala Ser Ala Gly Asn Ile
 35 40 45

Gly Glu Asp Gly Ile Leu Ser Cys Thr Phe Glu Pro Asp Ile Lys Leu
 50 55 60

Ser Asp Ile Val Ile Gln Trp Leu Lys Glu Gly Val Leu Gly Leu Val
 65 70 75 80

His Glu Phe Lys Glu Gly Lys Asp Glu Leu Ser Glu Gln Asp Glu Met
 85 90 95

Phe Arg Gly Arg Thr Ala Val Phe Ala Asp Gln Val Ile Val Gly Asn
 100 105 110

Ala Ser Leu Arg Leu Lys Asn Val Gln Leu Thr Asp Ala Gly Thr Tyr
 115 120 125

Lys Cys Tyr Ile Ile Thr Ser Lys Gly Lys Asn Ala Asn Leu Glu
 130 135 140

Tyr Lys Thr Gly Ala Phe Ser Met Pro Glu Val Asn Val Asp Tyr Asn
 145 150 155 160

Ala Ser Ser Glu Thr Leu Arg Cys Glu Ala Pro Arg Trp Phe Pro Gln
 165 170 175

Pro Thr Val Val Trp Ala Ser Gln Val Asp Gln Gly Ala Asn Phe Ser
 180 185 190

Glu Val Ser Asn Thr Ser Phe Glu Leu Asn Ser Glu Asn Val Thr Met
 195 200 205

Lys Val Val Ser Val Leu Tyr Asn Val Thr Ile Asn Asn Thr Tyr Ser
 210 215 220

Cys Met Ile Glu Asn Asp Ile Ala Lys Ala Thr Gly Asp Ile Lys Val
 225 230 235 240

Thr Glu Ser Glu Ile Lys Arg Arg Ser His Leu Gln Leu Leu Asn Ser
 245 250 255

Lys Ala Ser Leu Cys Val Ser Ser Phe Phe Ala Ile Ser Trp Ala Leu
 260 265 270

Leu Pro Leu Ser Pro Tyr Leu Met Leu Lys
 275 280

<211> 309

<212> PRT

<213> Homo sapiens

<400> 209

His Ala Ser Ala His Ala Ser Gly Arg Gln Arg Gln Leu His Ser Ala
5 10 15Ser Thr Gln Ile Arg Trp Glu Pro Ser Pro Ala Met Ala Ser Leu Gly
20 25 30Gln Ile Leu Phe Trp Ser Ile Ile Ser Ile Ile Ile Leu Ala Gly
35 40 45Ala Ile Ala Leu Ile Ile Gly Phe Gly Ile Ser Gly Arg His Ser Ile
50 55 60Thr Val Thr Thr Val Ala Ser Ala Gly Asn Ile Gly Glu Asp Gly Ile
65 70 75 80Leu Ser Cys Thr Phe Glu Pro Asp Ile Lys Leu Ser Asp Ile Val Ile
85 90 95Gln Trp Leu Lys Glu Gly Val Leu Gly Leu Val His Glu Phe Lys Glu
100 105 110Gly Lys Asp Glu Leu Ser Glu Gln Asp Glu Met Phe Arg Gly Arg Thr
115 120 125Ala Val Phe Ala Asp Gln Val Ile Val Gly Asn Ala Ser Leu Arg Leu
130 135 140Lys Asn Val Gln Leu Thr Asp Ala Gly Thr Tyr Lys Cys Tyr Ile Ile
145 150 155 160Thr Ser Lys Gly Lys Gly Asn Ala Asn Leu Glu Tyr Lys Thr Gly Ala
165 170 175Phe Ser Met Pro Glu Val Asn Val Asp Tyr Asn Ala Ser Ser Glu Thr
180 185 190Leu Arg Cys Glu Ala Pro Arg Trp Phe Pro Gln Pro Thr Val Val Trp
195 200 205Ala Ser Gln Val Asp Gln Gly Ala Asn Phe Ser Glu Val Ser Asn Thr
210 215 220Ser Phe Glu Leu Asn Ser Glu Asn Val Thr Met Lys Val Val Ser Val
225 230 235 240Leu Tyr Asn Val Thr Ile Asn Asn Thr Tyr Ser Cys Met Ile Glu Asn
245 250 255Asp Ile Ala Lys Ala Thr Gly Asp Ile Lys Val Thr Glu Ser Glu Ile
260 265 270Lys Arg Arg Ser His Leu Gln Leu Leu Asn Ser Lys Ala Ser Leu Cys
275 280 285

Val Ser Ser Phe Phe Ala Ile Ser Trp Ala Leu Leu Pro Leu Ser Pro
 290 295 300

Tyr Leu Met Leu Lys
 305

<210> 210
 <211> 742
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(742)
 <223> n=A,T,C or G

<400> 210
 cattgggtac gggcccccctc gagtcgacgt atcgataaagc ttgatatacg attcggcacg 60
 aggcccgacc gctccctgag agccagaac gggcagtgat gtttagcccc gagaaaaaat 120
 tacatgcgga atggaaagca ggcgctcagg gtggctctg ctggaatgag agctggagtg 180
 caggctccgt ggttccctggg catgcgggtg tggctcagtt ctcacctgc agatggagtg 240
 ggactgttga cccaggccag cctggggact gcctcctcac ctccctgcgc aggctgaccc 300
 tgcaccccttgc cctcttgcgc ttgcctctct cctgcccaga ngtccttgg acaaaatgg 360
 ggtcgagagg catttggcac tcacgcctca ccacggacac tggtgcatc ttgggtaccc 420
 ctggccctca atctattgtt gggggangga ngactgangc ccattgttgg ggccttgaat 480
 gcagggactg taaccaccca tcccctctc agggcacctc tccctctcca gcacncttgc 540
 tttgttattt atgttacttta atttcctact gangtggctt agaagcttctt ccgcatttgc 600
 ctttggccgc agcaaaatttt tatcccttagg gttaagataa cagaaggccan ctttgggcct 660
 tgccttccac atttcaggt ntncactgaa gcacagtatc tatttctcca aaaatagggg 720
 ctgttaactt gttactaccc cc 742

<210> 211
 <211> 946
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(946)
 <223> n=A,T,C or G

<400> 211
 ggcacaggc acatcgctgg atttctcatt gccaagctct attaattcat tcttttcat 60
 aaccttttat tcttatttca tggatgcaac attttcttgc tctctcaggg aataataatt 120
 attcctactt ttaaaaggctt aatttcttta ttactttatt tctctggag tgagtttttc 180
 ctaaaaggat aatgagatgg aaaatgaaaa aacaaagttt agacatggag ataccttctg 240
 aaactcaagc attcctctac gtggatgtgc cagggaaaa gaacagaaca aaggagggtt 300
 gacacttattt aaataaaaaat atataagaat attacataac aaacaaaaaa gcccaaattcc 360
 tcaggttggaa aaggaggaga aaatgtcaag caagacaaaa acagatgaag caaccaaaaa 420
 agtgacatag ctggtcaccc atattgaaat ttcaagacat gagtgataaa ggactcccc 480
 aaaaaaaacaa aacccaaact aaaaaacaga aaaaaaggac ttaccacccn aaaacttgan 540
 gaatcaggaa gactcagtct ctcattaaaga aaantgctat agggatggg ggcaaggcc 600
 tcaaagtngc agggatacc aataacctct ctgaagtttt ggaacttcat actccaaaaat 660
 ngttattttt tttttttttt cccgggtttagg ggccaattttt aggacttaga aaggaccng 720
 gnaatcatt cccnncttgc cccccccgaa agaaattaat agaagggtt tattcccgcc 780
 attannaaaa aaggaatcca ggaattncgg ntttttcca gtgttangnt gggntgtan 840

aaactgaggg cttagcaagg gcggnattaa ccacccnnggg tcccacccca aaantggnnng 900
gggtgggccc caaattcggg ntnttnccct ttaangcgtt aaaccc 946

<210> 212
<211> 610
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(610)
<223> n=A,T,C or G

<400> 212
ggcacgaggt ttctggctgg agcctcgac actggctcac tgcagtttgtt ggtgtcgaca 60
gtggtagag ggcaccagt aacgggagct tctcctgcctt ggcaggaaga cgagtagaaag 120
ggagcggcat gctggaggct ggagcctgag cccctggggc tcgccttgct gtgtttgggt 180
gtgacgtggg acactgcagc tcggccagag tggtaaaaaaa tgccttggtt tacgttttc 240
tggcttgcc cgtctatctg ctccaagcca ggctgganga ngagganaag gaatcacctg 300
tggtagcgtg gagcctgcat gtggcgtgac tctgcactc gcctcggtt actgatggca 360
gccacggaga ctgcagctg acagggagtg aggctctca ntggcttggaa agctcagctg 420
actcccacga aatttgcggg aaactcaagg ctgtcagtga ctttcgtggc gccaagactt 480
aancangcgc gttgcatgca tccggccagt gtcttgccat cgtgcctgca cnccacctt 540
anataancac ccggAACGCG cnncgcgcag gccgcgcga cacgnccggg cancaactt 600
gctggcttcc 610

<210> 213
<211> 438
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(438)
<223> n=A,T,C or G

<400> 213
ccganagcgg tttaaacggg ccctctagac tcgagcggcc gcccctttttt tttttttttt 60
aaataaattt ctagattatt tattacataa gcagaccact gaaacattt ttcaaaagta 120
ttccatttgag agtcaaaaac atattgatat gattattttt ggtctgtttaa agaaaaacaaa 180
ataaaaaagaa caaactggga attatcaata aacaaatcaa aacttagatg taattataac 240
ctaaaggcgt cacagggcaa atgtgaagca agcttctgtc tcagagccctg catatggaaag 300
acatgtatca ctttagcttttgcatcttttctt ttcctctctt tggttgagtt taagtattaa 360
taaaagggtgg actgagaaaaa ctttttttta caatctttagt gggtagttttaa agtggaaacg 420
ttttagaagt aggaatat 438

<210> 214
<211> 906
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(906)
<223> n=A,T,C or G

<400> 214
 gcccctctaga tcgngcggcc gcccctttttt tttttttttt gaaataaaatt tctagattat 60
 ttattacata agcagaccac taaaacattt attcaaaagt attccattga gagtcaaaaa 120
 catattgata tgattattat tggctgtta aagaaaacaa aataaaaaga acaaactggg 180
 aattatcaat aaacaaatca aaacttagat gtaattataa cctaaagggc tcacagggca 240
 aatgtgaagc aagcttctgt ctcagaccc gcatatggaa gacatgtagt acttagctt 300
 gncatcttgc ttccctccctc ttgnttgagt tttagtattaa taaaagttgg actgagaaaa 360
 cctttttta caatctttag ggttattttt agtggaaacg tttagaagta gaatatacat 420
 attaaaactg cncagaacaa atgnggtgca tctcaaattgg nngtccattt tcaaaaatgg 480
 aacacatag ggcagcattt ttttttttaa aaagtcagaa ggggcctnct catgccccctt 540
 tccacttctt cactcattgg nccttcaacc caagcttaac tactntcctg acctccaaca 600
 tcataaaacta gtttccnagc tttgaaactt ttttccaatg agtcntaccg gaatagatgn 660
 tcacagaanc ctcttaaaaaa ttttggaccc tgcccggnnt ntaaaaaggg tgcaataaac 720
 ccaccaacat ttggctggg ggggcagggg caaaaagaan ttcccaaaac cgtttttgat 780
 naaaaaaggg gacttttgaa aaaaaaatta aaattttgc cagnaaagca tgggnccccc 840
 cccttgaana aacccttgc atnaaaccaa ctttntgggaa ntttttngg tangttttt 900
 ctggct 906

<210> 215
 <211> 312
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(312)
 <223> n=A,T,C or G

<400> 215
 ggcacgagga aaccagggtt gctgggtttt ggggtgtaaac ttaaaaatga caatcagcat 60
 gagctggccg tgggctgtgg gggttgttagg ggcattttgg taagggaaacc ctcgctcagt 120
 ccctctctgt tctgggtgggg aggacaagga gggccaatag gggccaatag ggaggctgct 180
 gctaggangg tttcctaaaaa gaacaggtgt agggctaggg ctggttctt gttcaggtt 240
 ctctggccag tgatttataat ccacacaccc ttctgcaaaag tgcctaaagg aganggcagg 300
 gataggagtg tc 312

<210> 216
 <211> 341
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(341)
 <223> n=A,T,C or G

<400> 216
 taagcctntc gaanataatg aatgagtcattt ggagaggctn atgangaaat nccaaacacc 60
 tgactaatng gtgccacatg attncaatgg nctanacatg ggttagatct cttcngnngaa 120
 atgagcaata acaccnttaa antcncatg tgacccatg acttcacact tgaaanatca 180
 tcacttttaa ngaccacgaa tgatgcttaa gaatcacatt ttgtgnngaa ntggantctg 240
 gctacttaca cgaacagatt cttattccgt ttcatgagcc agtagaccgg gaanaagact 300
 taagagcttc tgancttct cttagctcca nngcttgaan g 341

<210> 217

<211> 273
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(273)
<223> n=A,T,C or G

<400> 217
nnccttcncc ccttnacnga catgaacaaa acagcngtct ngaaattttt ttaacatnn 60
aagggttacn ctccctnctt ntgtttccg nttaanncta nacctgcgcn ggggcggccg 120
atncagccct atagtgagaa gcctaattnc agcacactgg cggccgttac tanngnatcc 180
cgactcggta ncaantttt gngtaaagat ggacatanct ctatccnnga gnactcgtca 240
nccnttctct atnttacatg cnctaacgna gac 273

<210> 218
<211> 687
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(687)
<223> n=A,T,C or G

<400> 218
ttttcagtgc tggtttgttc tcaattttga tgtaaaaatc tctgggttct tctaantng 60
ttatgttctt ccanaaaatc cttccagttt ttgttaatttt tttctatatac agaagcgcct 120
gancccaatag cccaaatnat acaccggtct tctccggaac gcttggtna aagggtntag 180
tcnattnngc tcctggaagc atctnaaatg ctccaggta ctcccangnc cctggannac 240
ttcanttgtc tanacgaatc ctggttttcg agcggtcctt gatatcgcaaa ggaataacgg 300
taaaaattat ccaagctctc ttcccactna gganttcgga tctcatcagc cgggtaaagg 360
aaaactcttc angaagtttg ggcttccctt ccggtctacc ggctaattgtt aggaattact 420
tctggctctc ttccgataca tcctcttcc aaagtnaaga aggttaaaag aatnttaacn 480
tctcccagtg gctaatggtc aaacaccatc ctcatnagtc agactgggtt ttcgaaagga 540
ggatataacc tccttgcnag ttnnaattaa aagggattaa ccanatggac tancctcnc 600
cccgaggattt nctctctcac aggagaaggg gtctnccnc ttggctcatc cgaagcatag 660
gcaaaccnnn gggatatttc agaaacc 687

<210> 219
<211> 247
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(247)
<223> n=A,T,C or G

<400> 219
gggccttcn ccttnaatc gagagatcca aggttcaagg catgaaatac cagnctataa 60
aatgtctcaa gacntaaata atacggatng ngatagagag gttgaataat aaatgaanaa 120
anatgaaagn nattatnggg gaatacnaaa aaancngact aanggcggca ctgctggca 180
tggnnnaatc ggattaattc ctcataggac agccnaaccc cttaaaatct canttccgt 240
nacccga 247

<210> 220
<211> 937
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(937)
<223> n=A,T,C or G

<400> 220
cgggctcgag tgcggccgca agctttttt actatagacc aatattaaag tcagtttgt 60
tccaaataca ganttgaaa actaaagtaa aatatttaat gggagaatat ctgcattctga 120
atatgtcaac tgtttgcata ttttcagcta tttatccctt ctacctgtat ctcagaaaca 180
aattttaaaaa ttatagatt tgacagcaaa atcattcagc actttactta ctccatcagc 240
aaggatttta ttagtgcatt tccatccatg tggccaaact gaaaatccct aaccaccacc 300
aaccaaaaat aaataaataa aaggagaggg ggtggggggg gagagagaga gaaagctcat 360
taaatagtaa aaaagtaaat aaaacaatga agttaaattc aggcctcagt aggcggagaa 420
actgttaaca tttcacatgt aaatcatat caataaacac tgctaaaatgt gttaattctt 480
ctggcttcgt agataacaat acacgagtag agggaaatttct aagacatttc tacttggttt 540
atgcataatattt aaatttcagg gaaatatcag ctatttacc tgaaatatgt ttaagaaaaa 600
ttcctatttt ctctaaaaaaa aggaataatc agaagacgct acataactatg taagaaaaact 660
atacaatgac ccattcattag aagattcaga atagggaaaga aataataattt cactaataaa 720
atataatttat attgactgtc tttttttatg atagcaacaa tgattcagca taaagtaaaa 780
atataatgtat ttccgatgcc attttttattt cagttattct ttggatgtt ttttagaata 840
attatctgcc tatctctgac ttctgancag tcatttatgt ccaattataa gtacatgtgc 900
atattttattt accttaaacg cctctcaaattt ccttca 937

<210> 221
<211> 353
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(353)
<223> n=A,T,C or G

<400> 221
ggctatnnna tnnttnaan atcntgnncnn ctttgacgct gttantaaan aaaaacaaac 60
gaataatcctt ttttgcctcc cccctgtcna gataactatc tcacactaat acttacagta 120
taactnttcc ttctcaactac caatattaaag ttccaagcca cctgggctta agtatccaa 180
caacttaggt aattttgtgc taaccaccat actatatgtt aattataaca ctctaagccc 240
caaggaattt ttgttcagat ttcttattat ttccacttat aaattnattt ccncctctat 300
gggtattnn nnccctctagn cccattnnc ccacngggat ttgttgaggg ggc 353

<210> 222
<211> 813
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(813)

<223> n=A,T,C or G

<400> 222

ggcacgaggc ttactaaagg ccagactcac tatccccgt tctgttctgt ggtacactgt 60
 tcactccta gtcacatccta acctgacttc ctggccactg cagctctcc gataagggtc 120
 agcagtggct tagttattgc taaataataa ggcacatgc actcccttcc tcctgaaaca 180
 ttgtccctcc ttggtttctg ttcccttccca ggtctccat cactcccttct tagtcttctg 240
 tgcggacttc tggcccttcc gccctttaaa agtggatt ttccaggatt ctgtccctagg 300
 cccactact tctcattctg cacgttcttgc ttgatgatt ctatcacatc cctaacttct 360
 gctgcccagt atgcacttaa aattcccaa tctgtatatac tggatctggc ctgtgtctct 420
 agcctagaag tggctttat cccagaagca cctccaaacac tgcactttgg aaattaagct 480
 tactgagtct ctagtctcaaa gtcctttaact gacttcttt tctctatttt ggtagtgac 540
 aacactattt attcagtcattt gcaaaaccaga gccctgagaa ccatcttaca ttctttct 600
 cccttactc agttcttgc tctgttctt ctccctccncc tctctgcct gtggcctag 660
 ngnncattaa ctgggtggca ctgcttact ttcnatttt ttggctganc taaccnaag 720
 ancctntgt agggcctt ctntcaggcn tnacttctnn caagancccc cgaaccaga 780
 tccngggan tgctatggnn tggaaatattt ttg 813

<210> 223

<211> 882

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(882)

<223> n=A,T,C or G

<400> 223

tcacactact gagaaggcagg gaaacccact gaaaggcac gtttcttaac ctcagaatgg 60
 ggctactagc ctctaaagca ggaattgcgt tttttagt atttccatgg tctgctgcaa 120
 ggcgtggct ttacccaatg gataaatgcg tacaaggctc ttgtgagcag tcaagttct 180
 cgaggtttac agttgaaggg aagtggatt gtttccctgc gcatttaat gaaggtaggt 240
 gggtgcacatc ctcccttaa atgtgtgaag ggatgagata aagagatagg catcttaatt 300
 gcccactgtg gccttcaggt gaggacaggc atgagccaa tgaagcttgc acaattgtgc 360
 tgaacccaaa acttcaaaaa caagaaaaaa catagactgg ctgaaatgat ctaagtcaac 420
 agagcatggc cagcgcttca tacaaggcag gaccacaggg gaacactgac agccaggag 480
 gcactgagac agaggcagtg ggaagaagtg acagacccca gggactcccc accaacagca 540
 gctgctgttgc attaggaacc cccagtagac tgcaggcac ctggtagtgg agaggctacc 600
 aaggcccgga ctggagagga gccaaagaa gaaacagtgc agtgcttgc cccctctggg 660
 tctgcccgtg tccatcccccc tagggagatt ccattccaga agtggacata ttcccacaga 720
 gtgcctgggg ctcaactcatc acagctgccc ctnatgaag gcattctcac tgcagccctt 780
 ncagggaaaca gggtcatttg cattaggcan cttgctgtcc tagaaggcnt cgggnngtccc 840
 tacactgccc atgttcccaa ngnggttcaa nctcnaaaan tn 882

<210> 224

<211> 660

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(660)

<223> n=A,T,C or G

<400> 224

gattaaactc aatcattcac ccgggctcga gtgcggccgc aagttttt tttttttt 60
 tttttttt ttttggncct ctgggttgcg gcccggaaagg ggantgctgg gcccacntggg 120
 tgccgtgtt tgatttctg ggacctgccc ccccgtncc cgccccggnt gcccgtctc 180
 actccccgccc gccggcnaag gggccccgtg tgccgcgcac ccttccaccc gtgtttgt 240
 gttttttga ctntggcgt cccaggggtg cancggcgt gggggccctgg tttgcttca 300
 ccttttcatac tgctcaactgg cccgcancgt gncttnttca aacaaacgtt tgaaggncaa 360
 nccctggcgt cctgtgaacc cggccgtctt tgccgcaaaan tctgaggttc cttegttatt 420
 ctggatccgg ctntggcgt gggccgtgtt ctgcaggcac tgctccatt gctggcancc 480
 ttttttcccc gtggccgccc gggccgcctt naaggcgtt gcaaaccccc gcccctcgcca 540
 gcgcaaaagtc aaacncgggtt gggccgcgga ccccccggcg gncgggaaca cccccancagg 600
 cgggaccac aanaagcgcg gncctccggc gtctaaaact nccatgtggc ncccccccg 660

<210> 225
 <211> 438
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(438)
 <223> n=A,T,C or G

<400> 225
 aaaaaaaaaaag gaaaagtacc cagtgccttc agcttctgag cctcctctac agccctgttg 60
 gnntttaaac ctgtgccctg tgtctgtgtc cccacttaat atatatagtt cacagctgga 120
 gagatggctc agccaggaga gggaccata ggtctgtgaa ttccagagga naggcaggna 180
 tttataggtg gntctgtcag gtgaaatcng aggagccaaa gctattgtat gtgcataatgt 240
 cagccggcgt ctgtgggagg tggtgtaaga cctatggnat gggacangt tncacgctgg 300
 gatctctggc cggttccgaa aagtgaggat caggtgtgg gtggctgatt gcacaagttt 360
 anaaccagg attagggaca cacaggtcag caccgttcc tcagcatctt gactgggtgt 420
 gatggcata ctcaggc 438

<210> 226
 <211> 480
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(480)
 <223> n=A,T,C or G

<400> 226
 aaaataaaaa cccaaaggat ctttagaggc ctttacttca gtgggtctca atgtcagagg 60
 atgttatgtt acctaataa aatctccagg ggaactgttt tgaactcaac agactcttc 120
 ctgttctgag agactctggc aaagttggga gagctgccag gtactgtcca catgaccctg 180
 actgcccatg attcaattac cttgaatggc ttatccagtc caataccctt atttcttaca 240
 tgaggaaaact gaagcacgtt tcacatagt atacaatgaa aacttggcct taatcgattt 300
 tcagtgctgc cagttacaatg tcttgagcat atcaatttct tccaaaccctt gacaacataa 360
 ggtacgacca tcaaattttt tatttctgtt aatttattag accaaaaaaaaa aagggnatct 420
 cncccatgtt ttacaggga tgattttattt ncagaggatt tcacntgn gctgattcnt 480

<210> 227
 <211> 423
 <212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(423)

<223> n=A,T,C or G

<400> 227

cattgtgttg ggctctgttt agcacatcac atcggagcac agaggtgacc tgttctgcc 60
cagggatgtt caccttagtc acctgattga ttccctttca ctttggtcac gtgatttcctc 120
caggaggatg ttcaccttgg tcgcctgtt cctccaggag gatgttcacc ttgggtcgct 180
gaccacacag gcatctatca ggctttctca ctgcagccac tatgtccccca taatggatga 240
gtgtcttggg gagagatagt ccaaattgaca ctgataccctt ttgccttcata cggcctcacc 300
ccccacaataat cnaccactaa tgactgcctc atagcagttt ttccatttcc acagtttcctt 360
ctatatgtat taattgtcat tctactataa agaanacttt ttcttttaaa aaaaaaaaaa 420
aag 423

<210> 228

<211> 249

<212> DNA

<213> Homo sapiens

<400> 228

cattgtgttg ggctgttagta aaatatgtgt ctggtaagat atgtgaagaa ataaaaataag 60
atcaattaaa tctggcccat tgaatgacac attaattgtt tattaatatg taatgttaaa 120
gatatttagga gatgggtggga cattatggca aactaaattt gggaggaggt tgaattgtat 180
aatttatgaa atcctaaagt ctagtacatt aacactctct actgtcaact tttcaaagca 240
gtgagaaaac 249

<210> 229

<211> 436

<212> DNA

<213> Homo sapiens

<400> 229

cattgtgttg ggatgttatac tgaccatcac aatatgattt ataatatgga ggcataagt 60
catttctcat tggggcagga gtgtggcaag ggggaagaag agctttacca attaactcaa 120
gattatttgg tgacatttct cttacctttt aggtgaggag aaagagacag aggtggaga 180
atgggtgctt ttatgtatgtt gatacattaa gctgcctgga agcagatgtt aaatcctatt 240
gaaaataatt ttatgtcggt tttgctttagg gcattgttta gcaaaataact acacaaaaag 300
tcttgacctg tgggtttgaa atggcagatg ttcacagtga ggactgagcc ttggggcaac 360
atcaatcttc acaattctgc acctatttgc tcaataactg gcttgggtgg aaaaaaagg 420
aaaaaaaaaaa aaaaag 436

<210> 230

<211> 760

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(760)

<223> n=A,T,C or G

<400> 230
 cattgtgttggaa ggnngtggaa ggaaaanttt gaggcaatga agctaaacat aaaagaggaa 60
 aagcanatgt tacctaattt accacaatct acaaagtcca aatanaaaac ctgggagat 120
 gataggatga aactataacc tccagcaaag agcttaacag caattaaaat aaagacaaat 180
 ttctggatg gatnagacaa agtagcatat attacaagg aaaatanaact agtacatnt 240
 acgttgatt aagtaactgc tttcaataa ttgaatcata aacaatgatt tctgcgggtt 300
 taagctcatt atttggttc cctggttct ccttaggatgc agtataaat ctccatgcct 360
 gatgtttagt taccaacaga agctgctgct tcttcttc attatttcct ttttaagtga 420
 aagtaatac cttttatatg ttacagagaa gaggcagaaa aagccacact cccactatgc 480
 tattaaatgc cctgaggatc aactgaggaa tgattatacn catggctgaa tacagtnat 540
 tcattgttt ctggattt gatnataacaa aaggtgtat tctgttaacat cttgtgncaa 600
 ttanccaaat gtaaggcga aaatggaaatc tttcaacaaat gttntaaa caggtttga 660
 ttttccaaa ttantatta gaaccnttc aattctggaa gtnccaaat ttccangttg 720
 tgtttctct tccaatttctt cttcccttgc naattcccc 760

<210> 231
 <211> 692
 <212> DNA
 <213> Homo sapiens

 <220>
 <221> misc_feature
 <222> (1)...(692)
 <223> n=A,T,C or G

<400> 231
 cattgtgttggaa ggggggtctn tggggagaac acgttatgt tganatnggg ctccccgaga 60
 aagcctcatt gacacnttcg aataaggacc cttggggaaa ttcangtgag ttgtggacat 120
 ncntagataa natcaaaggc cttgangaaag tccgcctggc accttccngt ctgcgaggag 180
 gttgatcca aatgctaagg ggtccagntg cantgtanta tcgtgagatc agagtgtatgg 240
 gcaggtgtgg gcatgcgggc cctcaanang aagtgcggc gatgactcag acttatgcct 300
 atatccattc antcctgttc attatttttca ncntccctc naaggacccc caatttnaac 360
 cattgttat tcanggttat acttataaaa gtcattttgtt ttnagtcgg gtgatattaa 420
 aaccatttgg acgccangca tggtggtctn nggctataa tccntccac cttggggaaag 480
 ccgaagctgg tnnaatccct naaggtcngg aatttggaaa ccatcctggg ncaacattgg 540
 gngaaaccct gtctctactn caaaaaacan aaaattttctt ggggcctngg ttngcaggtt 600
 gcctgaaaat ttcccanctn tactccggaa agggcgaatg ccntaaaaaa nnaccttta 660
 acccccccga anggcggaa agttccatt tn 692

<210> 232
 <211> 518
 <212> DNA
 <213> Homo sapiens

 <220>
 <221> misc_feature
 <222> (1)...(518)
 <223> n=A,T,C or G

<400> 232
 actcaaatgn ccncttgaag gtcacccaga ctcanaangt gtcaagctt gggtggggtt 60
 gtaatnaata nctcggncctc ctgattagtn ctcctagtc gatcncctggc tgagatnngt 120
 tcgagcacccttcccttgcat cccgtcaaac nccngnnaaa agcngcctgc gtagtcncct 180
 nagccgaatc tgnnttcccg acacccctccg ctggctggc tgcctggtn aagcngcnc 240
 ctnaaanaan aaagngaagt ctccccngtc tcncccnant cctngggaaa acngcctgaa 300
 ccaatatgnt ccccccaaggn cnccccaggg cacntaaccg gtttaggaggc ccccccncntg 360

gcgtttggcn cnnaagcccn gccccngnnaa taaccccncct anaaccacgn aaaaatgcaa 420
agtcccaaag ggttaaagaat ctcnnaccc cccggttccc tcgcaanctt cccctnngna 480
cttgggttcc gggaaaaaccc ttancccgan ctttcca 518

<210> 233
<211> 698
<212> DNA
<213> *Homo sapiens*

```
<220>
<221> misc_feature
<222> (1)...(698)
<223> n=A,T,C or G
```

<400> 233

```

gcacgaggtt ctgtctgtct gtctctctct ctctctctct ctctctctgt ctctctctca 60
cagttagaat ttggctgttt tcttttattca ataccccaat atatgttcat tagggttata 120
ctgtatacac tacacataac agttttgttt tttgttttgg atattatttgg ataataagaa 180
tttaccacca tcattaaaaaa aagtttcccc aagctataat ttttgataat tgcacttcc 240
cactattcaa atgaaaaat aactctttct ctccctggagt aggtttacat tccattttag 300
ctatgataact gctttaagag aaattgtttt aagataaaatt tccatagaca ggtcaagga 360
ggtaatata tgtaagctt tcgatgcctg ttactgaatc tcattctggaa aacataact 420
gtcaatgccc tctttttctc atggaaaaaa aatacataac aaaatttacc atcttaatcg 480
tttttaaatg ttacagtacg atagtgtna ctgtatgtac ttgtgcaac agattctctg 540
aaaactttt cattttcaa aatgaaaaact ctgtactcat tgaacaggca gcttccaaac 600
ttccccatc ctcccannc ctacccttgg ttaanagtct nacaaaaccc ggaaatttta 660
tggaaatttga aacacttttta naataccncn tattaggg 698

```

<210> 234
<211> 773
<212> DNA
<213> *Homo sapiens*

```
<220>
<221> misc_feature
<222> (1)...(773)
<223> n=A,T,C or G
```

```

<400> 234
ggcacgagcg cagctttcg aaagctgtaa ttgtttgt atcaaaaagtc ctgcagttata 60
ttatgtctcat tgcattttaa agagttcca agtgatcagt gatggttgtc tgtttttag 120
tattacggtc ttatgttaatg ttgcggaaact agtcagtttgc tgctgtcgt acggggccgaa 180
aagatcaggc caggccaaagt actctggccg cccaaagtaaa tgcttaaggc cgccaaacggaa 240
ttatgtccctg ggggttcgatg agggccgtaa ttaggtttag ctgggtgtang ctaaccttcgc 300
agccatgtcg gagagagatg agagacataa nattttaaag taggggctgtat ttttacgaag 360
ttctgancca tttcccttgc tatcggtccc ggcaaaaagca actgagataa atgtgttaaa 420
agactcgatg attttttcgaa cttcagcaac gtactcagcc ttgggttctc gtagtttttc 480
aaaggcagct atttgctgag attcatgaaa agtttgactt ganctgcttg tcaattttctg 540
cagcncgggc ttcaactgtt attgaatttg ttgtatataag cncaatacgt tgcnngtccac 600
caagggtttc catgtttga ctncacctgg tcgaaccaat ttgaattatg tntttttgc 660
tgnccctgttc cccncctt aaatccatct cttttttngaa aacctttngng nngttgaattt 720
cnqccqccqccq gttcccaacn tttqgttca ccttqaaaaaa aaanatqqt aqgt 773

```

<210> 235
<211> 849

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(849)
<223> n=A,T,C or G

<400> 235
attgggtacg ggcccccctc gagcagcctc cactgcaatg ccgctgaatc aagagacttt 60
tcaatacgt ttatcagtga aatgtatgtg atctgaagag tcctatcttgc agcactttgc 120
atgacatcca acgttaatgt ccacaacgtt cttagctgcc caacccctt atcggcaagc 180
tccaaagggtg tgtcaaagc ttctacggcg tcatgaaaag ctgaaaatg ctgtgtcaac 240
actgcacccgc tgcgcacatctt caaaagcagc gcccttatacg tctccgcatt cgaagacgt 300
aaccggcgtt gaatagcctc ataatactt ttgttagaaat caatcagacg tgtgttagga 360
acctttccat caaaaacata cgactgtgcg accacgtctg caaaagcaga cgtcacatta 420
tgcataatgcc ctcttaccgt cagccgatca tcctcaactca tagcgacgcg agaaagctct 480
tgttccagct cgtgcacggt atccaattca gtaatcttac gcaacgcgtt ctgaatgt 540
ttcataagtt cagttttaaa gctcaaaact tcgtcttta nttaaaaaat ttgtgactttc 600
aaactggcg antcttcacc attttattaa tcgtctttt gangganggc ccagcgttag 660
atctgcacgc ccagcggaaat cgttactccc tccttacgtt cctccggta acgcanntag 720
tttctccgaa gccttaaaat tagccgggaa aagggaantt atttgccccca acaanggnat 780
cgccggcctg gtggttaaaaa ggaactgaaa taaaattaaa nccccncttgg gggaaangcc 840
cgcataactg 849

<210> 236
<211> 310
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(310)
<223> n=A,T,C or G

<400> 236
gggggtgggtt gcttccgaaa nccggggccc ggccaacttg ttggcttggg aatattctgg 60
caagaaaatt tccagggcgg cgccaaatttcaatggccctt aaccggaaaac 120
tctggcagggtt tcaacccctt tcatgggcgn ttggaaagctt gaagcgcccc aagttactcc 180
caagcttggt gcgnttgcgg ttggggggcg gggaaaagtt gaaaacacgg gcgnttggtt 240
gccccccccc cgggcgggttt ntacgcctt cctggggaaa ctttcagggtt tggctgttta 300
cnaaaacggg 310

<210> 237
<211> 315
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(315)
<223> n=A,T,C or G

<400> 237
gcacgagttt ttgttattta natnttgctt tggttaangg aagaacacaa naatgcctg 60
ctaaaggat tctgtttggc tgcangctgc nacggggaa aaaatcnaan tgtatnttgc 120

acaacangat ttttagaan tcagaactat gacatgaagt canncagggc actctacgac 180
 tgaatttgcn gtgctgcctt cacangctcc ttntcgctc tntnctggca ncngtgactc 240
 ntacacgtcc tgganantan cctccctana aggaacgact ccgacacccc cccnntaccc 300
 ctnaangttc atcng 315

<210> 238
 <211> 510
 <212> DNA
 <213> Homo sapiens
 <220>
 <221> misc_feature
 <222> (1)...(510)
 <223> n=A,T,C or G
 <400> 238
 ngcacagtn ttgttattt atatattgtt ttgtttaaag gaagaacaca aaaatgcct 60
 gctaaaggga ttctgtttgg ttgcagggtg cnngcggggaa aaaaatcaaa gtgtatttt 120
 cagaaaaatga ttttttanaa gtcagaacta tgacatgaag tcaagcaggg cactctagga 180
 ctgaatttgc tggctgcct tcatatgcct cttgctcgct cttttctggc agctgtgact 240
 cncacaggtc atggaganta tcattcccta aaaggaacaa cnccgatatt catcttata 300
 cattaaatgc atctgtccca ttctatgtng tggatgctaa cttttgatca ttgatngtga 360
 tnccatggac atntancatc anctttcana ncctnggatc tttgacnagt cttattantn 420
 agantccaac tantacgatg ccganttana aatgctggnt ntccaattcc tactcaaata 480
 nccnacatga acttccantc cccttgcnnna 510

<210> 239
 <211> 209
 <212> DNA
 <213> Homo sapiens
 <400> 239
 ggtgcttttc ctttctactc gtcttcctgc ctggcaggag aagctccgc tactgggtc 60
 ctttctacca ctgtcgacac caccaactgc agtgagccag tggccggc tccagccaga 120
 aacaggtagc accatgccg gataccaaac gcccacactt aagagcctga aatgacctga 180
 cggcacctcc gcatgttta cctactgag 209

<210> 240
 <211> 610
 <212> DNA
 <213> Homo sapiens
 <220>
 <221> misc_feature
 <222> (1)...(610)
 <223> n=A,T,C or G
 <400> 240
 ggcacgagggt ttctggctgg agcctcgac actggctcac tgcagtttgtt ggtgtcgaca 60
 gtggtagag ggcacccagt aacggggact tcttcctgcac ggcaggaaaga cgagtagaaag 120
 ggagcggcat gctggaggct ggagcctgag cccctggggc tcgccttgcgt gtgtttgggt 180
 gtgacgtggg acactgcacgc tcggccagag tggtaaaaaa tggccctgggtg tacgttttc 240
 tggcttgcc cgtctatctg ctccaaagcca ggctggganga ngagganaag gaatcacctg 300
 tggtacgctg gagcctgcac gtggcgtgac tctgcaactc gcctcgtgtg actgatggca 360
 gcccacggaga ctgcagctcg acagggagtg aggcttctca ntggcttggaa agctcagctg 420

actcccacga aatttgcggg aaactcaagg ctgtcagtga ctttcgtggc gccaagactt 480
 aancangcgc gttgcatgca tccggccagt gtctgtgccca cgtgcctga cnccacctt 540
 anataancac ccggaacgcg cnncgcgcag gccgcgcgca cacgnccggg cancaactt 600
 gctggcttcc 610

<210> 241
 <211> 474
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(474)
 <223> n=A,T,C or G

<400> 241
 ggcacgaggt ttctggctgg agcctcggac actggctcac tgcagtttgt ggtgtcgaca 60
 gtggtagag ggcaaccaat aacggggagct ttcctgtcca ggcaggaaga cgantagaan 120
 ggancggcat gctgganget ggancctgan cccctggggc tcccttgctg tgtttggtgg 180
 tgacgtggga cactgcagct cggccagant ggtaaaaatg tcctggtgta cgctttctg 240
 gcttggcccg tctatctgct ccaaggccacg ctggaagang agganaagga ntcacctgtg 300
 gtacgcccga gcctgcatgt gggngtgact ctgeaactcg cctcgtgtga ctgatggcac 360
 ccacggacac tgccactcta cagngaatga ggcttctccn tggactngaa agctcanctt 420
 nactccncc aagtttgncc gaactcaagg ctntcactna acttcgtggc gcca 474

<210> 242
 <211> 415
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(415)
 <223> n=A,T,C or G

<400> 242
 ngcgggnnt tccaccagct cgtgtgcaca agtngcgcca cacaaacatg cgcaaggact 60
 gcatgtcata natgtgttcc gccgtggttc tggAACAGCG agtagaagat ggcgttcggg 120
 tcgcgaccaa attcgacgtc ntggatgttc ttgcgcaga aangtcacgtt cgggatcgac 180
 ccgatggatc cgctnaagcg ccgaaaggcc ctgacttgca aaccgcggct cacagaacccg 240
 gcaccacccg ccccttcggc cnacaaaatg cgagcggctt ccgacacacaca ctccctcaca 300
 tccccgtcnc gcaatttcggc nttttcttagc tccgcacgg ttgtcagcgg caccgcgggc 360
 gcccagctgc cggcggcata cgttgcacac agcacacacg gatccgctct cgtgc 415

<210> 243
 <211> 841
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(841)
 <223> n=A,T,C or G

<400> 243

aacgaggtgt cgatgagcgc gaacaatcgc cctccttcat ctctacctga tggtaactt 60
 cgctcttaca gccgagccaa tgaagacgaa tggctgctgc cgaggatgg agtctcaact 120
 gagcacgcgg cgctggacaa ctcatcgact ttgtacgctt cggtagctt gcccattcag 180
 ctccactgac gacagagacg gagctggca ctggccatctc gacgcagcgg gacaaggagc 240
 agttcgggc gccgtatgca tcactcgaag agaaccagga gcagccggaa gcaggangcg 300
 ctgcacggta caggcacttt cggcgcttca gcgatccat cggccgatc cctgtacgtca 360
 ccttcttgcg caagaacatc caggacgtcg aattcggctcg cgaaccgaat gccatcttct 420
 actcgcttcc ccaggaccgg gcgaaagcaca ttgtacatc gcaatgcctt gcgatgttt 480
 gtgcggcgtt accttggtgc acacgaacga nggcaaccaa cccgccccag gtgcgcgtct 540
 atgcattctt gttctgttcc ggtgtgcattt gccggatgtg gacgtganc ttggtaatc 600
 ggctgggtca tgaagactta cccgtcttca aagggcgaa cgcncctcan ttgganaag 660
 gaacaaaacc ccccccnaag aacggcattt gcanctttt ccccegtgc cggcttctt 720
 ccattcgggn attctctntc tccnaaaantt cccgnaaaatc ttcttgcgtt ttctccctg 780
 tttttatgg cccctcccgca cacttgggtt gtttacatc ctacaancct ttttttctc 840
 c 841

<210> 244
 <211> 761
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(761)
 <223> n=A,T,C or G

<400> 244
 aacgaggtgt cgatgagcgc gaacaatcgc cctccttcat ctctacctga tggtaactt 60
 cgctcttaca gccgagccaa tgaagacgaa tggctgctgc cgaggatgg gagtctcaact 120
 agagcacgcgg cgctggacaa actcatcgac ttgtacgctt cggtagctt agccattca 180
 gctccactga cgacagagac ggagctggcc actgcccattc cggcgcagcgg ggacaaggag 240
 cancttcggg cccgtatgc atcactcgaag gagaaccagg agcagccgaa agcaggagc 300
 gctgcacggta acaggcactt tcggcgctt cggatccaa tggccgat cccgtacgtc 360
 actttcttgc gcaagaaaca tccaggacgt cgaattcggt cggcaccggaa atgcacatctt 420
 ctactcgctc ttccaggacc cggcgaagca catttgcgtt actgcagtgc ctgcgcattgt 480
 ttgttgcggc gtcacctgtt tgcacncgan cggcaac aacccgcgcc angttgcgc 540
 tctatgcatt ccctgtctgt cgggtttgc atggccggat tggancgtg ancttgc 600
 tccgctgggt gcatgaagga cttaccgctc tcgtcaaggaa cgaacgcgcc atcaattccg 660
 gaaaaggaaac naaaaaccccc ccccaangac ggnnaatttgc anctttccc ncncctgc 720
 gctcttctcc antncggct tctttctc anaaaattcc c 761

<210> 245
 <211> 710
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(710)
 <223> n=A,T,C or G

<400> 245
 aacgaggtgt cgatgagcgc gaacaatcgc cctccttcat ctctacctga tggtaactt 60
 cgctcttaca gccgagccaa tgaagacgaa tggctgctgc cgaggatgg gagtctcaact 120
 agagcacgcgg cgctggacaa actcatcgac ttgtacgctt cggtagctt agccattca 180
 gctccactga cgacagagac ggagctggcc actgcccattc cggcgcagcgg ggacaaggag 240

cagcttcggg cgccgtatgc atcactcgaa gagaaccagg agcagccgga agcaggaggc 300
 gtcgcacggt acaggcactt tcggcgcttc aegggatcca tcgggcggat cccgtacgtc 360
 accttcttc gcaagaacat ccaggacgtc aaattcggtc gcgacgaaat gccatcttct 420
 actcgcttt ccaggaaccc gcaagcaca ttgataacat catgcctgcc catgtttgtt 480
 gcgccctcc tggttgcnca cgaancgaag gcaacaaac cccgcgcagg tngccgctct 540
 tatgcattcc ttgtctgttc cggtnntgca tgccccggan ntggaaaccc tnancttgg 600
 nnaatcggtt ggtgcattga aggaacttac cgtctcgtc aaggccgaa cgcnccttc 660
 agttcggana aaggancgaa aaccccccna aaggaacgg ccnttgcnng 710

<210> 246
 <211> 704
 <212> DNA
 <213> Homo sapiens

 <220>
 <221> misc_feature
 <222> (1)...(704)
 <223> n=A,T,C or G

<400> 246
 aacgagggtt cgtatgagcgc gaacaatcgcc cctccattcat ctctacactga tggtaactt 60
 cgctcctaca gcccggccaa tgaanacgaa ntggctgtcg ccgaggatgg gagtctca 120
 aaagcacgcg ggcgtggaca actcatcgac ttgtacgttt ccggtagctt agccattca 180
 gtcactga cgacaganac ggagctggcc actgcctatct cgacgcagcg ggacaaggga 240
 gcagcttgg ggcgtatgc catactcgaa agagaacagg agcagccgga agcaggaggc 300
 gtcggccgtt acaggactt tcggcgcttc anggatcca tcggccgat cccgtacgtc 360
 accttcttgc gcaanaacat ccaggacgtc gaattcggtc ggcacccgaa ttgcattctt 420
 ctactcgctc ttccaggggac cggcgaagca cattgatnaa attgcattgc ctgcgeatgt 480
 ttgtcgaaaa cttctgtgtt ccccgancga agggcnacaa ccccgccca gggtgcnc 540
 ctatgcattc ctntctgttc cggtgttgcn tggccggat ttgaaccgtg aancttgg 600
 aatccgntt gtcattaaag aacntaaccg ttntcgtca gggcnnacc ggnccctnc 660
 aatttcggaa aaangaacca aaanccccna ccnccaagga aacn 704

<210> 247
 <211> 618
 <212> DNA
 <213> Homo sapiens

 <220>
 <221> misc_feature
 <222> (1)...(618)
 <223> n=A,T,C or G

<400> 247
 ggcccccagt gtatggata tcgaattcaa cgagggtgtcg atgagcgcga acaatcgccc 60
 tccttcatct ctacctgtat gtgaacttgc ctcctacagc cgagccaaatg aagacgaagt 120
 ggctgtgcc gaggatggaa gtctcaactgc agcagcgcgc gctggacaac tcatcgactt 180
 gtacgttcc ggtagcttag cccattcagc tccactgacg acagagacgg agctggccac 240
 tgccatctcg acgcagcggg acaaggagca gcttcggcg ccgtatgcat cactcgaa 300
 gaaccaggaa gcagccgaa gcaggagcg ctgcacggta caggcaactt cggcgcttca 360
 gcggatccat cggccgcata ccgtacgtca cttcttgcg caagaacatc caggacgtcg 420
 aattcggatcg cgacccgaat gccatcttctt actcgctttt ccaggaccgg gcgaaagc 480
 attgtatgaca tgcgtgtcct ggcgtt gtcgttgcg tacctgggtgc acacgagcga 540
 nggcaacaaa cccgcgcaca ggtgcgcgtc tatgcattcc tttctgtcc gggtgtgc 600
 gggccggatg tggaaaccc 618

<210> 248
<211> 622
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(622)
<223> n=A,T,C or G

<400> 248

gcacgagagc ggatccgtgt gtgctgtgtg caacggatgc cgccggcagc ttggcgcccg 60
cggtgccgtc gacaaccgtg gcggagctag aaactgccga aagtgcgcac ggggatgtga 120
gggagtgtgt gtcggaggcc gctcgacttt tggggcgga gggcgccgtt ggtgcccgtt 180
ctgtgaccgc cggtttgaa gtcagggcct ttcggcgctt cagcggatcc atcggccga 240
tcccgtacgt gaccccttgc cgcaagagca tccacnacgt cgaatttgggt cgcgaaccga 300
acgcatctt ctactcgctc ttccagaacc cggcgaagca cattgacaac atgcnnntgcc 360
tgcgcattt tgcgggcgc tncctgntgc acacgaccga ggttaccaac ccgcggcagg 420
ntgcnnctct acgcatttctt gtctgcccgg tgcgtggc cnggatgtgg accntgagcn 480
ggngantccg ctggtgcnng aagacnttgc cgctctcgctc aaggccnacc gcccnnctcg 540
gcggaaaaaag gancaaaaanc ccccccggccaa gaaccggcnc tgcaccgttn tgcgcggccct 600
gctgggctct tcccnnttac gg 622

<210> 249
<211> 517
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(517)
<223> n=A,T,C or G

<400> 249

cattcgagct cggtaccggg gatccgattt gtaaaggaaa tgcggaaacag ccagctgggt 60
tttcgggtgc ggcggggca gcccacatcg ctgtgggtgt tggcgtaactg gatgcgtatgt 120
gcggggacaa acgcgttttc caccacgtg tcatgactgc ctgtgcccgg caggcccagg 180
acatcccagt tgcctcaat gcggtagttcc gccttggggca ccagaaaaagt cacatgtcc 240
aggccaggcg tgccatcaag cttggggcagc agaccgccta gaaacagcca gtcgcaatgc 300
ttggagccgg tggaaaaagct ccagcgaccc ttgaacctga atccgccttc cacgggctcg 360
gccttgcag taggcatata ggtcgaggcg atgcgcacgc cgttatcctt gccccacaca 420
tccctgtggg cctggtcggg gaaaaancgc cagctgccaa ggggtgaacg ccgaccaccc 480
cgtaaatcca ggcgtggac atgcagccct ttaccaa 517

<210> 250
<211> 215
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(215)
<223> n=A,T,C or G

<400> 250
nntncattgg gccgacgtcg catgctcccg gccgcccattgg ccgcgggatt accgcttgtg 60
accgcttgtg accgcttgtg accgcttgtg accgcttgtg accgcttgtg accgcttgtg 120
accgcttgtg accgcttgtg accgcttgtg accgcttgtg accgcttgtg accgcttgtg 180
accgcttgtg acnnggggggtg tctgggggac tatga 215

<210> 251
<211> 231
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(231)
<223> n=A,T,C or G

<400> 251
ngcggccacc tngtgattga tggtcgttta ctatcaagta tgtacatctt gctctagaca 60
actccnattc agtggaaag attgggaaag tatcccgat aagtaatagg nattaggtct 120
nccttantgc ttgggtggat attccncaac tgnntccngat cggatcagnc tcgtgtcngn 180
gaatgtgctc gatcgtnatt ctactnctga gcttctatcc nnacgtggcc t 231

<210> 252
<211> 389
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(389)
<223> n=A,T,C or G

<400> 252
atgtatcanc nctgttggtg ttncatctt tgcagtcngt tctaaggcn gataantatc 60
agagatgcta atgcatttc tgccaggcca ncattggtgg cctatgcgtt ctcttctt 120
cttcctgaag agtcatctt ggnngatgtg ttccccccctc tccacagtgt ttgcaagcgt 180
taccacgcn tgcggngcc gggaaaggtn ncacatccgg gnagacttcc ccncgtntga 240
atcgtnctn gaatctccgg cgtcncctt naaccttctg actnggacaa ngnccctnt 300
tccctntgt gaactngtan cccgccttcc ttccccccctc agcctaancg ggaangaaga 360
cngggtnat ctnnggcncc acaagaant 389

<210> 253
<211> 289
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(289)
<223> n=A,T,C or G

<400> 253
nggggcnna ttagcgcgta taatacnatc actatngggc gaattggta cgggcccccc 60
tcnagcggcc gcctttttt ntttttttt ntntttttt caaaaacaccc tccnccntgg 120
atgganacgt naccttctc taaccanatc ttcacaatnc nantctcagg cagccgcctc 180

aaanccgatg tcangttggn atntcaantn caatcttatt ttgngaatta anctganatt 240
gtggatggtn naccaatcan atacttggn a tccgttgaac ccctgtgga 289

<210> 254
<211> 410
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(410)
<223> n=A,T,C or G

<400> 254
attgtgttgg gaactttag acagctata caattgcagt gctatttctc tgaggtattg 60
aatctcantt attataattt taaaatccaa ttggcttggaa cttcatttatt ttccaaactaa 120
aaagatgattt gaaggatttta tttgaaatgt gtaaagagta atatagattt tatgtttatg 180
tttccttggaa aaaagtaggtt aaaatttctc tggaaagtgtt actcctaaaa tacaaatgaa 240
catgtcaaga attacataaa ttctttaaac tattcttaan aannaatggc tctatgtann 300
gagngaccct tacagactat taagaattaa ctgcatttgc anagactcat ttanattcat 360
gaaatggntc tcaactttctt ggttaagatct ggcttggacg tttttggtaa 410

<210> 255
<211> 668
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(668)
<223> n=A,T,C or G

<400> 255
ttttttttt tttcctgtg ccaggcacta taccactgtg cttaggtgcct tcttgcatt 60
acttcatttc ctcataagct ttctgaggan acagaaagct tgaggttcac gtagctagca 120
tctacataaa tttagttgcta aaaacataca atacgttcc cggcaggctg tcatttagtaa 180
ctgataactac tagttgataa ttcataaaac ttagcanaan ctaccattta agctgaaaca 240
actgtcaata tcaactaanta aaacttaaat ccataaatca actatattct aaaatctgac 300
ttcaggtaa ttaaaaaatc actagttgtt acctacctcc ttctgaaagc cagtacaagt 360
taaatgaaca actcccgagt ttaacaaaca agtggcatct aaaaaaaaaaaga tttaaaaat 420
aatccactta catatattta aaatggcatt aaaaaacaa aatttatcca ataacnaant 480
ggcaaaggaa ggtgtccaat tattacatgt tataaatctt taaattaaac ttttcttngg 540
ttttctntcc ctanaataaa tacaancctt tcccccgcna accagaaaaa agcaaaaaac 600
aaaacccaaa aactcccagc ncngcttaaa aaacnccaaa aaaataaaa ctctttaaaa 660
tgcccnnaa 668

<210> 256
<211> 487
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(487)
<223> n=A,T,C or G

<400> 256
 cgnaaccgtn cnttttnat gtgcgcccgc cncagnacca gngccgctac aggccaaggc 60
 cggaagcacg ggagaggntt ngaaaaaaa agagtgtta caaagagcat attcgcagag 120
 ttgggatgag tgaaggggac cagaaggngc acggtaggg acgcgtgaaa ggangcngcg 180
 gagaaatgac agcaagaagg gganaagcac acgaaaaggc agtacccctt 240
 tcgaggactg ccgcacatctt gtttctgcc cattccagtc accgaanaag atccaaana 300
 aagaagaaaa gaancagagg tgcacttcgc ttcatatttc nctcgcttc tttctgnct 360
 tcacnagttc tgcaggattt cccttgcctt cttccgagca catctacga cgnatgaggg 420
 tcggcaggc aagccnacaa aacnctcgca ctccctttt tcttgcnn tctgngtgg 480
 angnggg 487

<210> 257
 <211> 502
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(502)
 <223> n=A,T,C or G

<400> 257
 ccttggaaag nccngctnaa ttcngnganc cccngatca gcaccaggga gctacaacna 60
 aggccgaaag caggggattt ngccggaaaa aaaagagtgc ttacaaagag nttatccnca 120
 nagatggat gagtgaaggg gacgagaagg tgcaagcgta gggacgcgtg aaaggaggca 180
 gcggagaaat gacagcaaga aggggagaag cacacgaaaa ggcagtatcc tcctcccccc 240
 ttttcgagga ctgcccgcatttttctt gcccattccca gtcaccgaaa aagatcccaa 300
 agaaagaaga aaagaaacag aggtgcattt cgccatcatat ttgcgtcgct ttctttctg 360
 tcttcacaag tctgcaggat tgccctgtc ctctccgag cacatctacg cacgtatgag 420
 gctcggaggn caagccaaaa aaacgcttgc actccctttt ttcttgcgt gtctgtgtgt 480
 atgtgaaattt ccgcggcncc gc 502

<210> 258
 <211> 510
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(510)
 <223> n=A,T,C or G

<400> 258
 actcgncaact cgatncanta caagagnnta tgnattcgaa ngtgcccccg catcagcacc 60
 agggagctac aacgaaggcc ggaagcaggg gagagggccg gaaaaaaaaa agtgcttaca 120
 aagagcatat ccgcagagtt gggatgagtg aaggggacga gaaggtgcag cggtagggac 180
 gcgtaaaagg aggcagcgga gaaatgacag caagaagggg agaagcacac gaaaaggcag 240
 tattcccttc ccccccttcc gaggactgcc gcattttgt tttctgcctt ttccagtccac 300
 cgaaaaagat cccaaagaaa gaanaaaaga aacagagggtg cacttcgtt catatttcgc 360
 tcgctttttt ttctgtcttc caagtctgca ggattgcct tgcacttcgc cgagcacatc 420
 tacgcacgtt tgaagctcggtt aggtcnngnc aaaaaaacgc ttgcacttctt 480
 gcnagtctgt gtgcatttnggg gaaatnctna 510

<210> 259

<211> 292
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(292)
<223> n=A,T,C or G

<400> 259
gannngagtc acgaaaaggc agtatacctcc tcccccccttt tcgaggactg ccgcacatcttt 60
gttttctgcc cattccagtc accgaaaaag atccaaaga aagaagaaaa gaaacagagg 120
tgcacttcgc ttcatatttc gctcgcttc tttctgtct tcacaagtc gcaggattgc 180
ccttgccttc ttccgagcac atctacgcac gtatgaggtc cgaggatcaa gccaaaaaaaa 240
cgcttgcact cctttttttc ttgcgtgtc tggatgtatg tg 292

<210> 260
<211> 582
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(582)
<223> n=A,T,C or G

<400> 260
gcacgagggtt ggggtgtact gtgtataata actccagatc cttgaccaag tttggagagt 60
caacctatggc catttgaaac caaatgaagg atcaaaggac taatttattt gaataacctct 120
gagtgttttc cccaagcttg agaagagttt cattcagcta taaaatgctc atttgcaaa 180
tgagtgtttt ccatgctgtta taattaaagc attgccttta ataataatttt attaccttta 240
gcttgccttt ttaatttgag gaaaatccaa acaatttaaa gtaaaacgtg ataaagacag 300
tttttcngga gananaaggg nagatcgcta tgtttattcc acttaatatac tataatcaa 360
atttgatca aaagcagact ctcactttaa aaatattttt ctaatggcna gaatctttt 420
ccttagattga gagtcagagc tcacatagna tnactgctgg taaatagaca ctttagactat 480
agagctnagc tnaagttcca actanccaaac tgcatgttg aatatgctt ttattnaaag 540
gccagnnctt ttgcctttt nccnccctaa tnccttctat tg 582

<210> 261
<211> 783
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(783)
<223> n=A,T,C or G

<400> 261
gcacgaggca aaatacagag ggtatttac catggacagg caacccattt ttccaggaca 60
actcttgcgca gcagagagct attctctttc ttttgccctta cactctcaac ctcactttc 120
gagtgtctgc atcctanttt tccatggcca taagataagg aaccatgaat gttactctag 180
atgaggctgt ttcatgtgg gagctcatcc aggatccaa gtagattcat cagaaggta 240
agtataggag tggaaaccca aatctctact ttatatttga ggccttctt cctcaatttt 300
aaattgtaaa atcaaactta aaactggta tctgatggcc agttaaaaga ctgggtatct 360
gattgccagt taagagatgg tcatttatgc tcaccaccat tctcaagacg caggtgaggt 420
gacangctt ctggggaaatg ctgancgaat ccccaatgc cttcaggatt ctggaaatgg 480

tggctctgnt taaaactgn tgaactttac aaagagccta cccgtcatgg ggggactggg 540
 aagaaaaccc anangcagt tctggccan ggttacaccc ccangntac cttgaaggnt 600
 tttggacat acctnttncc cccctttac tgnttcatta gggcncnnc aacccaant 660
 tccaagttnt ggccttcna aaanttttt ntttccnnt tccanggacc cccctggntt 720
 cctggnnccc ccttttata nccaaccttg ccnggnattt ttcncnntn aaaggaaat 780
 aat 783

<210> 262
 <211> 741
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(741)
 <223> n=A,T,C or G

<400> 262
 tgaaccctan tggggccggc cccctcgagt cgacggtatac gataagcttg atatcgaatt 60
 cggcacgagt gtatattctg ttattatacc ccagattnaa gtgtatattt ttaggcagta 120
 gttctggta acatccttac tacataaaaat ccacttacta tttaagtattt attctaacag 180
 gaggtagaat agctgcctta aaaaatgttag tgatcgaatg gcagttttc tgctgaatgg 240
 aaattactga cacaatttattt ggttttggga gacattttcc tccttgggt tgagttttcc 300
 cattcacgga tagggcataa agcttgggtt atagttgagg ggtgcaaaag gggaatagga 360
 ttgggaaaat acagtgttcc agcaaaggta tgacaaggta catcttgag aggattccct 420
 ttctgctang tggcactgtt aactttttttt tactgtgtac ttccagaca aaggatagag 480
 aaaaagacct tcactgggtt ggggagaaga aaacccttgc tccttagaaaa atcacaaaaa 540
 aggcatcctt tancctataat tcccagntt actggngcat ttgcttgatg tgactgacnc 600
 ngattatttc cttnactgg naaaaattcc tgccncttg gatatnaang ggggnaccng 660
 gaaaatnggg ggcnttgggg aaggaaanaa aaaaattgg agggaccnnaa ctttggaaaa 720
 tgggntgctt nangccttaa g 741

<210> 263
 <211> 437
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(437)
 <223> n=A,T,C or G

<400> 263
 ggcacagag aatgtgttca cagacactat ttatanaat tctgtatgtt actgtgtctg 60
 gtggatgtga aagccatact tcttaaatct gatttggaaa gcaaatctga ttatcacagc 120
 cataattaaa ttggccagc ctcccttctt ccctccctcc ttcaacttct tccttccttc 180
 cgcctcgtgc cgaattccggc acgagcctga cctcaactacc aaaaaaaaaa aaattcaag 240
 tgcctgaggt ttccaggcat tcttagctt atttacttac ttcccaccc tcataatggcctt 300
 agaattcaaa ttctgnanaa aatggattgc catanataat ccaatggaaa tgggtcatat 360
 ttggccatta atagaatcac agtcnacaag ggactaatag aatttagtcaat ttagtatacn 420
 ttagatttgg gagacnn 437

<210> 264
 <211> 706
 <212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(706)

<223> n=A,T,C or G

<400> 264

gcacgagcac cccaagggtt taggacaaaa tggatgagt gaattcatgg cttgacagac 60
 tgaacagaaa aatgagggctc cgtgctccat attcatgtgc atctgccct catggtaca 120
 tgctaattgg ttggccgggtg cacaagacaa ggaagtgcag gtttctgtt gtcacacag 180
 tgcttcctgt ctgctgtggc aggagccggg aggaagggag cgagccaaga ggggtgtgc 240
 ccacccggaaa cgatggcgcg aggccgaga gctaaatggg ggcctctcca gggagtgtc 300
 tgttcacggc tccatcgtcgt ttagtaagta tcttgtgatt tcggaattta aatgagggtt 360
 tgtttaacct gcataacatc tggctttaa aatctgactt tattttctt ttatttctgt 420
 gcatggctc aggcacactt agtggggct tagtgttga agtcaggta ccaaacagca 480
 cgcctctct ttatttctcag gctgcgtgtt tcattgattc tgaagggtcag atggctgtgt 540
 tcaagttctg tttagtatatt ggtgtcagaa atgaaaagat gatgtaaacc tttataactt 600
 cttaaaggct catatcatgt cagggaaatc acctgtacga gttatggaca aatgcccattc 660
 ctgatgattt tcancatga aatgaatna aaggganaa gggcca 706

<210> 265

<211> 717

<212> DNA

<213> Homo sapiens

<400> 265

ggcacgagca gcattacggt ttatacacat gtccacaact cagcattgct ttcaaaatag 60
 gaacacttta ttagtaaaa ggaagaaaatt gcctaaacag actcagtgtc tttccataa 120
 caatcatctg ccaagccgca ggcctaacca ggaatccc tttccctttg gctgtgtc 180
 ctccaccaac agatacaacc ctgatgcca atgttgtatg gttttaggt gttgtgagcc 240
 aatgagggca tgccttagggc caaaggctgc ccttggat gagggcaagg tcgtagactc 300
 catcaaacaa caaatgcac ctcctccaaa atcaaatgct caacacatgc agccttcgt 360
 atgcccattt cccctttact cattttcatg gctgaaaatc atcaggatgg gcatttgc 420
 ataactccta caggttaatt tcctgacatg atatgagcct ttaagaagtt ataaagggtt 480
 acatcatctt ttcatttctg acaccaatc actaacagaa cttgaacaca gccatctgac 540
 cttcagaatc aatgaaaacac gcagcctgag aataaagaga gggcgtgtc tttggtaacc 600
 tgacttcaac acctaagcca ccactaagtg tgctgagcc gatgcacaga aataaaagga 660
 aaataaagtc agattttaaa aagccagatg ttatgcaggg taaacacaac ctcatta 717

<210> 266

<211> 362

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(362)

<223> n=A,T,C or G

<400> 266

ggcacgaggt tagatttaac ttccacagat gactcagcag aggataacta ctaatcagag 60
 tacaacatca aaactgtaac cagtataatc actggattat gagcaactca aaatagctcc 120
 agtttccaaa gggccataaa ctgcacatcat cagtaatgt tgcaattta acataattta 180
 ttatgaaaat gtggacatgc caggtaaatc aggggattt ggttacttt ttataatact 240
 ttaaatttga aatgcattt ctgtggattt gatgcacatct tccaggtgtc ntaatnctgg 300

gntacctnct gatanatcct gananaaaaga ggtancacca gcgtctatca nacctcaata 360
ca 362

<210> 267
<211> 692
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(692)
<223> n=A,T,C or G

<400> 267
ggcacgaggt tagatccaac ttccacagat gactcagcag aggataacta ctaatcagag 60
tacaacatca aaactgtaac cagtataatc actggattat gagcaactca aaatagctcc 120
agtttccaaa gggccataac tggccctttt aanaactttnn gcaattaaca cataatttat 180
tatgaaaatg tggacatgcc aggtaaatg ggggatttag gtgacttttataataactt 240
taaatttcaa atgcccatttc tgggattgg atgacatctt ccaggtgctt taatttgggtt 300
tacctcctga tagatcctga cagaaagagg nacgaccaggc gtctatcaaa cctcaataca 360
gngtgtgaaa cacangagag cctgcttttgcnacacggg gaaacacatt gttatcacaa 420
cacacaaaag gcaanctncc aatgggnan ncttacactgn cctctcatat tggggcaan 480
gaaaangggg ccccccatttgcgatgatn tcccaaaaaa ccnccactan tggtcagnnt 540
gcttccccan acagccatgactgatattt agcccaagct gcaatctcaa aaccagctt 600
ctgacaatca gtaacaagaa catactggtc tggcagtg agtcaagtg ttgggtgttc 660
agtcaaaaanc catggatgcc aatcatctcc ca 692

<210> 268
<211> 605
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(605)
<223> n=A,T,C or G

<400> 268
cgtgcgaat tcggcacgag ngcacatatc agtactatgt gcaattaaca cataatttat 60
tatgaaaatg tggacatgcc aggtaaatg ggggatttan gttgacttttataataactt 120
taaatttcaa atgcccatttc tgggattgg atgacatctt ccaggtgctt taatttgggtt 180
tacctcctga tagatcctga cagaaagagg tagcaccaggc gtctatcaaa cctcaataca 240
gttggaaaac acagagagcc tgcttgccata cacatggaga aacattgttacacaagaca 300
cagaaggcaa acttccaaatc tggcataactt ncttgcctc tcataattgg ggcaatgaga 360
atggggacc agatggctt antagatgcc aaagaacacc canactgggc agcatgctt 420
cccagacgc cngaaagactg aaatttanc ccagctgcag ncttaaaccc ttttttgac 480
nttccgtaac cagaccatac tttttttct gatgctttc ttaacttcat cttttccaaat 540
taaattcatt agtnnaaccc taaangggc ccgtttccg aaaaatttcc ntntntttt 600
ccccn 605

<210> 269
<211> 535
<212> DNA
<213> Homo sapiens

```

<220>
<221> misc_feature
<222> (1)...(535)
<223> n=A,T,C or G

<400> 269
gcacgagng caacccagg gtgggtctc tggatgaac ctggagacct gagcttgcac 60
agcttccttg gtaaattgag gaggcatgga ccacaagatt gccaagctcc ttctatcca 120
aacttgatatt tgtagattc catgatccag ttcacacgg ttgatggctg aatctcatgc 180
actanaaaaaa gtaatataa aaganaaaaa tanaangatn ttcaagttag tataaanacc 240
tttaatctca ntctttctag ttcaaagaga cggaaacaatg agagatgtg gttcatanag 300
ctgntanatt taacttccac agatgactca ncagaggata actactaate anagtacaac 360
atcaaaaactg taaccagtat aatcaactgga ttatgagcaa ctcaaaaatag ctccagttc 420
caaaggccca taaactgcca tatcaantac tatgtccat taacccataa ttattatgaa 480
aatgtggac atgcccangtn agtaagggga tttaggggtga ctttttatna tactt 535

<210> 270
<211> 803
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(803)
<223> n=A,T,C or G

<400> 270
gcacgagggc aaccccaggg tgggtctct gggatgaacc tggagacctg agcttgcaca 60
gcttccttg taaattgagg aggcatggac cacaagattt ccaagctct ttctatccaa 120
aacttgatatt tgtagattcc atgatccagt tcatcacgg ttagggctga atctcatgca 180
ctagaaaaag gtaatataaa agaaaaaaat aaaaagatat tcaagttagt ataaagacct 240
ttaatctcg tctttcttagt tcaaagagac ggaacaatga gagatgtgg ttcatalogc 300
tgtagattt aacttccaca gatgactcag cagaggataa ctactaatca gagtacaaca 360
tcaaaaactgt aaccagtata atcaactggat tatgagcaac tcaaaaatagc tccagttcc 420
aaaggccat aaactgcaca tatcagtact atgtcaatt aacacataat ttattatgaa 480
aatgtggaca tggccaggtaa gtaagggat ttaggttgac ttttataat actttaaatt 540
tgaatgcca ttctgtgga ttggatgaca tcttccaggt gctttaattt ggttacctc 600
ctgatagatc ctgacagaaa gaggttagcac cagcgtctat caaacctcaa tacagttgta 660
aaacacagag agcctgnntt gcctacncat ggagaacatt gttatcacaa gacacagaag 720
ggaacttcca tctggctact tacctggctt tattttggg gcaatganaa tnggggacc 780
aatggntgan tanatgccaa aaa 803

<210> 271
<211> 836
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(836)
<223> n=A,T,C or G

<400> 271
gcacgagggc aaccccaggg tgggtctct gggatgaacc tggagacctg agcttgcaca 60
gcttccttg taaattgagg aggcatggac cacaagattt ccaagctct ttctatccaa 120
aacttgatatt tgtagattcc atgatccagt tcatcacgg ttagggctga atctcatgca 180

```

ctagaaaaag gtaatataaa agaaaaaaat aaaaagatat tcaagttagt ataaaagacct 240
 ttaatcttag tctttcttagt tcaaagagac ggaacaatga gagatgctgg ttcataagac 300
 tggtagattt aacttccaca gatgactcg cagaggataa ctactaatca gaggataaca 360
 tcaaaaactgt aaccagtata atcactggat tatgagcaac tcaaaaatgc tccagttcc 420
 aaaggccat aaactgcaca tattcgtact atgtcaatt aacacataat ttattatgaa 480
 aatgtggaca tgccaggtaa gtaagggat ttaggttgcac ttttataat actttaaatt 540
 tgaaatgcca tttctgtgga ttggatgaca tcttccagggt gctttaattt ggtttacctc 600
 ctgatagatc ctgacagaaa gangtagcac cagcgtctat caaacctcaa tacagttgt 660
 aaacacagag agcctgctt gnctacacat ggagaaaacat tgtatcacaa gacacagnaa 720
 ggcaacttcc atctggata ctacgttct ctcttatttgg ggcatganat ggggacaatg 780
 ntgananatg caanacacca atngagctg ntccnacag cnatatgatt ntccat 836

<210> 272
 <211> 203
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(203)
 <223> n=A,T,C or G

<400> 272
 ggagaattgg gcccgtcang ggtgcattct gcatcacctg anttcnaaat ctnagtcaat 60
 cncgtacta atantatcaa catnattta acctgatctc cactgctng tnatttcnn 120
 ttcactgncc ctntcaactng aacntctntt cacacagcca ccccccattt tctggntggc 180
 acctccncca aatnccnccct naa 203

<210> 273
 <211> 594
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(594)
 <223> n=A,T,C or G

<400> 273
 attcggcccn ctggatncgt gctcgagcgg ccggcgtgt gatggatatc tgcanattc 60
 ggcttctgga gagagctttt ttttctgatgg ttgcangtac tctcgatgga gttgggtgggt 120
 gtggtatct ctctctgggt gtctttctgt ataaanttct tgcnctgact ncctanctcn 180
 cctccccctg gtccctccct tagngtaaca nctggtaatc cctntcttct ttgctctccct 240
 tncttctctt gancgttcc ctctntttgt ccactctcg gnanaaccct gntggtcagt 300
 gttcatgact tcnngaagnt cgacccgcna aatagggnen cacggatnat gttgaancng 360
 ggaagggagn gtcacaattt tctgttccan aggttnagcc tagaganaat gatgggagan 420
 ggtttactga gatcatngnn tcttctgaa gatatnntt agggtggccc cccataagng 480
 aatttctcan ctccaaatct tctaatacat tactgaacan ctgnatttgc ttacgccaca 540
 nattgnaatt ctccatntct ttttagaaac nattncaagg tcatttattt ccct 594

<210> 274
 <211> 229
 <212> DNA
 <213> Homo sapiens

<220>

<221> misc_feature
<222> (1)...(229)
<223> n=A,T,C or G

<400> 274
ctactactg tccggccatt tggncctctg natgcatnct caagcagcnc gccantatga 60
tnnatatctg cacanttcag cttctngaga aaactatgtt ttaaacagtt gcntanactt 120
anaatanaaa tcgagtaagg tntagatnan tctctaacga tngaattatt ntacanaggg 180
gtannncgatn accaggagta nctagantt ancancancc tagtcnga 229

<210> 275
<211> 651
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(651)
<223> n=A,T,C or G

<400> 275
atatctgntg aatacggntt cctgnaaaaa ggtntnattt agatggttga gtccgactca 60
gcgatgcac ttggtggtg tggtcantct cttatggttg agattgttca tgatatcatg 120
ccctgagatg cctggactnn cctcaccgga gatccttagac ggtgntancc cctgagagtc 180
tctctcntcc tgctctccta acttctccta atgatccctc cnattgtcta ctgtccnatt 240
gaacccttct tgcttatgtt tncaatcnn nacgggttcc ctgctnattt tttganacga 300
ngctcataat ggacngggga aggatagtnt gaataatntc ctgtataaccc acgcncnacnt 360
ctacnctntg atctgacacg gtatactgat ttgtgctgtt cncttcacca ttccantttc 420
tacctccgc tcatatgctc tgtangctac accctctgtg actgcttct cagttacgtg 480
caacaaggtn ttcatatctn gaactcttac accattctag anggatcncc cctccgganaa 540
antttggaan aacaaggcaag ancanaatnc ctctctngtg ntacacnanc cggcttncgt 600
atcctcgtn aaggaattcc ccgctttctt gggctttaan tctcctaaac t 651

<210> 276
<211> 392
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(392)
<223> n=A,T,C or G

<400> 276
accccccgg aattacgntg gcncatntaa aagtnccatca ngcctccang caacntatcn 60
tttcatttacc acccacactc ctgttnnggg anggangtgg naatccttca ccatnctaat 120
gtatgtggtg ctctcatgcn ggtacgtata atctannctg cccctnaat cggatgcttc 180
tgtatcnnc agtcacnnaaa ccacanggan caactgaaac angatttggc taacagccaa 240
tgtctgggcc ctncncaatc cctnnaatat ctcctacacc tgttagtanna atnaactacn 300
ctacnctatt nnacacacgn tttaggtgt annaccaagc ccntattgag tgaaatcggtt 360
tntatngtat naaatgccaa aagntgcggaa aa 392

<210> 277
<211> 212
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(212)
<223> n=A,T,C or G

<400> 277
ggttgcggg natgaanttt gnaanaatna acttagtnga taacccaccc accaatncct 60
nctnagtatt tgncaacctn aaaactacag ctctctccag atagactntn cctnctgat 120
ttcaactctc cttggactgg tcagcctgaa gggtggtaat gactcaccaa cgctactaat 180
ncctnttta ctgtgccttn atttttcgc ct 212

<210> 278
<211> 269
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(269)
<223> n=A,T,C or G

<400> 278
nnntccatcc taataccact cactatcggg ctgcgaancgg ccggccgggc acgtntcttn 60
tgngacagga tctgaatnaa gggtggtttg taacttnact naaaattctg aaatgatcct 120
gcatcagaca gggttctccg tntanaatan agttccctg ttagttatcn agcctggca 180
ggggangana gattcgagga cntntgaaat gaaggnatta ttttaggatgg gtgactcatt 240
ccnaccnttc ncgctnacca gnccganga 269

<210> 279
<211> 266
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(266)
<223> n=A,T,C or G

<400> 279
gttggtgant cngtttggng tcttccttgt gntnggtgtt tggtgtgttg nnttggtn 60
gggtngtnnt tntggagaga gttgttgtc gtgagggttg cagtgtactt actatggagc 120
ctaaggangt gngctaactt anantgatna ctttgctcat actgcccgtc cctnaatgcc 180
nngcttcctt caccctggtg ccnaaccnna tcaaacacctt aacagtctag taggcttctt 240
gctntancag actntcttgtt aggatc 266

<210> 280
<211> 317
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(317)
<223> n=A,T,C or G

<400> 280
aactgttag gtgtntggaa ntgntgtagg catagnctt ntggcacaga gttggagccg 60
tgaggcatag cngtactta ctatggagcc taaggangga gctaacttat antnatnact 120
ttgctcatac tgccctgcctc tnaatgccta ngcttgcctc accctgntgc cttacnnnat 180
cgaacaccta cgcggctat aggttcttg ctctatcagg actnctttc nagcttcntc 240
gcctcantg actcaactgtg ctgggtcggt ctactngat ccagncgctc atnaacctna 300
cttnggacgc aggtcat 317

<210> 281
<211> 174
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(174)
<223> n=A,T,C or G

<400> 281
gnggtcatat tatacatcta aggcattggcc aactccacgc cattatnaat tccatcgta 60
tgtccgcagt cactacttat aacctagatt aatagtgcct ggccccggac ngtctgtgca 120
atctnccgcc ataccaattn cgatccncan accncgatna cactcctcct tact 174

<210> 282
<211> 169
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(169)
<223> n=A,T,C or G

<400> 282
atcgcaagtt gtacgatcgt catataacgc gcatgtgcgg atcgcttcag cgccgcccga 60
ctgtcagaag gangagatct ttttatcac ttgtttgtt gactatanat aanancgact 120
acagcattga tgtgtgtcct caaganttg ctgggtctga naaagctga 169

<210> 283
<211> 157
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(157)
<223> n=A,T,C or G

<400> 283
ggntntctaa gatecgagtt gtacgatcgt catatnacgc gcatgtgcgn atcgcttcac 60
gtcgccngc tgccaggan atgcatntca acataatgtg cactctataat ggttattgat 120
taatacgagn tangagcana tatcngatac aacacaa 157

<210> 284
<211> 133
<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(133)

<223> n=A,T,C or G

<400> 284

ggngtgggt nagatacgca ngctgggacg aatcgntca tagtacggcg catgtgtga 60
tcaattctga aaatccatcc cggcgcgctc ancatgact anagggcaat cgcctatatg 120
antcgatata caa 133

<210> 285

<211> 194

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(194)

<223> n=A,T,C or G

<400> 285

ntntngtga tgatacccaa gctggntacc nactngantc caattaccgg ctcantntgc 60
tngaaacngc ttcgatngnc tcctggcatg tacttgaac agntanata tctaatagn 120
tacngtgnn tttcnatca tacagntnt atattncact ncctnccatt ctttctant 180
ctctctctcc ntat 194

<210> 286

<211> 134

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(134)

<223> n=A,T,C or G

<400> 286

gagggnntat gataccaagc tggtagcanc ccgtcaactat nacggccag tgtgtggatc 60
cgctanctgg tcncgcgatg tctacncaca cgngaactgc ctctcgcnna gatctcctct 120
cctctccnna gaga 134

<210> 287

<211> 119

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(119)

<223> n=A,T,C or G

<400> 287

tnggtatat ccagttgtac actggncata tacgcgcatt atgatcggtt cacgccccga 60
gtacggcata attacganat ggnctcatc gttaccttt ntcgctggac acaagcgtc 119

<210> 288

<211> 170
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(170)
<223> n=A,T,C or G

<400> 288
gggntgagat acncaagttg gtacgagtcg gatcatatna cggncgcat tttctggaat 60
ccgcttacgt ggtcccgcg aagtactttt tcattgccttg caaaatngcg ttactgcact 120
ancttgctta acctatgagt ggggtcttc ataccccntc tntcatggaa 170

<210> 289
<211> 126
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(126)
<223> n=A,T,C or G

<400> 289
ggccaattgg ggcctctana tgcntgctcg aacgggcgcc aatttnatgg atatctccaa 60
aattcggctt accntggtcg cggncnaagt acttaactca atccatctnt cactcaggat 120
naatgc 126

<210> 290
<211> 126
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(126)
<223> n=A,T,C or G

<400> 290
ggccaattgg ggcctctana tgcntgctcg aacgggcgcc aatttnatgg atatctccaa 60
aattcggctt accntggtcg cggncnaagt acttaactca atccatctnt cactcaggat 120
naatgc 126

THIS PAGE BLANK (USPTO)